

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 03-24-04
Searcher: Beverly C. 2528
Terminal time: 201
Elapsed time: _____
CPU time: _____
Total time: 28
Number of Searches: _____
Number of Databases: 1

Search Site
____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other CGN

[illegible]


```

/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAD18870.1"
/db_xref="GI:4377036"
/translation="MSFLIKPFSPPLKQDYLFDISPYTSSEIITGGSYFKLNKASL
QSSTLRLSISIIIS"
8654..9535
/gene="nfo"
/note="synonym: CPN0732"
8654..9535
/gene="nfo"
/codon_start=1
/transl_table=11
/product="Endonuclease IV"
/protein_id="AAD18871.1"
/db_xref="GI:4377037"
/translation="MKVLPPPPPLLLGAHTSTAGLKNAYIEGRDIGASTVQIPTANQ
ROMORRAKKEVIEDPKAALKETDLSYIMSHAGYLINCPADPVILEKSRIGIYQELI
DLITIGISPNHPHGAALKSKKEDCKNKLIVSFSOSAPLFDSSPLVLLVETTAGQGT
LIGSFEELGIDVQNLKQDPIGVCDVITCHIPAAGYDITSPQGMEDVINEFDEYVGLS
YLRAFHLNDSMPPLGANDRHAPLGEVIGKESFKELMTDBTRKIPYLETPGPGPEN
WQKEIGELKFKSKNRDS"
complement(9654..10283)
/gene="rs4"
/note="synonym: CPN0733"
complement(9654..10283)
/gene="rs4"
/codon_start=1
/transl_table=11
/product="S4 Ribosomal Protein"
/protein_id="AAD18872.1"
/db_xref="GI:4377038"
/translation="WARYCGPKNRVARFPGANIFGRSNPLLKKEPPHPPQGHGWRKKX
SDYGLQLEEKQKACYGMINEKQLVKAFKEVYHKGQNVQWFLRPFRCRLDNMVMYRM
GFKATIFAAQQLVAHGHLVNGRRVDRRSFFLRPGWQISLKESKRLQSVKDALBSKD
ESSLPYSISLDKTFKGLLVSPQDQIEAQLPLFINISVVCPEFLSHRT"

```

ORIGIN

```

Query Match      100.0%; Score 2156; DB 1; Length 10421;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1  ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTGAGAGACTTTTATTTATTTT 60
DB 5133 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTGAGAGACTTTTATTTATTTT 5074

QY 61  APAATAAACTAAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
DB 5073 ATAATAAACTAAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 5014

QY 121 AGGTCTATAGACGAAACAGACGACACTCCCGCAGATCTTCTGCTCAAGGATTGGA 180
DB 5013 AGGTCTATAGACGAAACAGACGACACTCCCGCAGATCTTCTGCTCAAGGATTGGA 4954

QY 181 GCGAGTGTCAGCAATAAGAGTGCAGAGCTCAAGAGANTAGCAGTGCAGAGCTAAGCC 240
DB 4953 GCGAGTGTCAGCAANTAGAGTGCAGAGCTCAAGAGANTAGCAGTGCAGAGCTAAGCC 4894

QY 241 TAAAGAACTAAGACCGGATTTCTGAGCGATGGAGCATCTTGCCTTCGAGTGAATGC 300
DB 4893 TAAAGAACTAAGACCGGATTTCTGAGCGATGGAGCATCTTGCCTTCGAGTGAATGC 4834

QY 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTAGTAAACAGTGTCTTCTACTAG 360
DB 4833 TCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTAGTAAACAGTGTCTTCTACTAG 4774

QY 361 CAGATCTGACAGCTGGAATCTCAACGACGACGACGACCTAGCCTCCCTCCACCCAGGTT 420
DB 4773 CAGATCTGACAGCTGGAATCTCAACGACGACGACGACCTAGCCTCCCTCCACCCAGGTT 4714

QY 421 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 480
DB 421 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 480

```

```

DB 4713 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 4654
QY 481 AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCTACTAATAATAAGGATAC 540
DB 4653 AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCTACTAATAATAAGGATAC 4594
QY 541 AGCGGCTACTGATGAGGAAACCGCAANTGCTGCGAGTGGGAAACTAAGAAATGCGGATGC 600
DB 4593 AGCGGCTACTGATGAGGAAACCGCAANTGCTGCGAGTGGGAAACTAAGAAATGCGGATGC 4534
QY 601 AGTTAAAGTTTGGCGCGCAATTAACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 660
DB 4533 AGTTAAAGTTTGGCGCGCAATTAACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 4474
QY 661 TCTTGACTCTTTAGGTAAACTGACTTCTCTTCGACTCTTTACAGGCTGCTTCTTCCAAATC 720
DB 4473 TCTTGACTCTTTAGGTAAACTGACTTCTCTTCGACTCTTTACAGGCTGCTTCTTCCAAATC 4414
QY 721 TGTAGCAACAATAACCAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATACCCAGTAGT 780
DB 4413 TGTAGCAACAATAACCAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATACCCAGTAGT 4354
QY 781 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGAC 840
DB 4353 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGAC 4294
QY 841 ACAGATAGAGAAAGATGGAANTGCAATTTAGGATGCAATTTTTCAGGACAGAGACGCTAG 900
DB 4293 ACAGATAGAGAAAGATGGAANTGCAATTTAGGATGCAATTTTTCAGGACAGAGACGCTAG 4234
QY 901 TGGAGCTCTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTCAGCTTAAGC 960
DB 4233 TGGAGCTCTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTCAGCTTAAGC 4174
QY 961 AGCAATCGCTACTGCTAAGACACAAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
DB 4173 AGCAATCGCTACTGCTAAGACACAAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 4114
QY 1021 AATTCTTCAAGAACGGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATACAA 1080
DB 4113 AATTCTTCAAGAACGGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATACAA 4054
QY 1081 ACCTGACAGATGGTCTGATGTTCCAAATCCAGGAACATACAGTTCGAGGCTCCCAAGCAACA 1140
DB 4053 ACCTGACAGATGGTCTGATGTTCCAAATCCAGGAACATACAGTTCGAGGCTCCCAAGCAACA 3994
QY 1141 AGGAAGTAGTATTGGTAGTATTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
DB 3993 AGGAAGTAGTATTGGTAGTATTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 3934
QY 1201 CGCTTCCATTTTGTATGTTGCTGGTTTGGTCAGATGATTCATGTTCAATAGCGAAATGCC 1260
DB 3933 CGCTTCCATTTTGTATGTTGCTGGTTTGGTCAGATGATTCATGTTCAATAGCGAAATGCC 3874
QY 1261 TGATTCTCAAGCTCCCAACAGGAGCTCGCAGCACAAGCTAGACAGCAAGGCGCTGG 1320
DB 3873 TGATTCTCAAGCTCCCAACAGGAGCTCGCAGCACAAGCTAGACAGCAAGGCGCTGG 3814
QY 1321 AGATGACAGTCTGCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGG 1380
DB 3813 AGATGACAGTCTGCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGG 3754
QY 1381 TAAAGCTGGGCAACAAACAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
DB 3753 TAAAGCTGGGCAACAAACAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 3694
QY 1441 TGTGAGCGCAGAGGTTCTCTCCGCTGCGAGCAAGTTCTATAGGTCATCTGTPAAAACAGCT 1500
DB 3693 TGTGAGCGCAGAGGTTCTCTCCGCTGCGAGCAAGTTCTATAGGTCATCTGTPAAAACAGCT 3634
QY 1501 TTACAAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGA 1560
DB 3633 TTACAAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGA 3574

```

1561 TGCTTAAATCCATCAATGATGCTATGCTAGGACGAATGATGCGACTCGTATGT 1620
Db TGCTTAAATCCATCAATGATGCTATGCTAGGACGAATGATGCGACTCGTATGT 3514
Qy GATTAACATGTAAGTACCCCGCTCTACACGATCCGCTTCTAGACGACGACGAGC 1680
Db GATTAACATGTAAGTACCCCGCTCTACACGATCCGCTTCTAGACGACGACGAGC 3454
Qy TCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC 1740
Db TCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC 3394
Qy TCTTGGAGATCTCTATAGTCAAGTTTCGCGACTCAATCTCTGAATGCAGATCATCCAGTC 1800
Db TCTTGGAGATCTCTATAGTCAAGTTTCGCGACTCAATCTCTGAATGCAGATCATCCAGTC 3334
Qy GAATCTTCAAGGATATGAGGATCAGACAAAGCTTACATCGGAGTGCAGAGCC 1860
Db GAATCTTCAAGGATATGAGGATCAGACAAAGCTTACATCGGAGTGCAGAGCC 3274
Qy TCCACAGTTTGGCTATCTTATGTCGAATCTTCTATGACTCTACACAGAGTTCAATAGC 1920
Db TCCACAGTTTGGCTATCTTATGTCGAATCTTCTATGACTCTACACAGAGTTCAATAGC 3214
Qy TAAATGAAGATTTGTTGCTGAAGATCTAGGACAGAGTGAATATAAGACATTTTC 1980
Db TAAATGAAGATTTGTTGCTGAAGATCTAGGACAGAGTGAATATAAGACATTTTC 3154
Qy CTTTGAAGACGACTCTGTTTATTCAGCAGTCTGCTCAATATCGGCTCTCTATATTC 2040
Db CTTTGAAGACGACTCTGTTTATTCAGCAGTCTGCTCAATATCGGCTCTCTATATTC 3094
Qy TCGTTATCTCCAATAACACACTTCTGTTTGGAGATTTATTTATGCTTTGGT 2100
Db TCGTTATCTCCAATAACACACTTCTGTTTGGAGATTTATTTATGCTTTGGT 3034
Qy AAGGCTTTGTTAGGCTTACCAACACACTAGACGATTTCAATAATAAAGA 2156
Db AAGGCTTTGTTAGGCTTACCAACACACTAGACGATTTCAATAATAAAGA 2978

RESULT 3
LOCUS AE002165 12548 bp DNA linear BCT 30-MAY-2000
DEFINITION Chlamydia pneumoniae AR39, section 2 of 94 of the complete genome.
ACCESSION AE002165 AE002161
VERSION AE002165.1 GI:7188948
KEYWORDS Chlamydia pneumoniae AR39
SOURCE Chlamydia pneumoniae AR39
ORGANISM Bacteria; Chlamydiales; Chlamydiales; Chlamydiales; Chlamydiales.
REFERENCE 1 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Unyam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, M., Craven, B., Bowman, C., Dodson, R., Gwin, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
TITLE Genome sequences of Chlamydia trachomatis Moyn and Chlamydia pneumoniae AR39
JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)
MEDLINE 20150255
PUBMED 10684935
REFERENCE 2 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Unyam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, M., Craven, B., Bowman, C., Dodson, R., Gwin, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

FEATURES
source

Location/Qualifiers
1. 12548
/organism="Chlamydia pneumoniae AR39"
/mol_type="genomic DNA"
/strain="AR39"
/db_xref="taxon:115711"
/notes="synonym: Chlamydia pneumoniae AR39"
297. .965
/gene="CP0011"
297. .965
/gene="CP0011"
/note="similar to SP:P31218 GB:X71492 PID:296947 GB:U00096
PID:1736770; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="uridine kinase"
/protein_id="AA037907.1"
/db_xref="GI:7188950"
/translation="MLMMLMIIGITGGGAGKTLTQNIKEIFGEDVSVICQNYIK
DRSHYTFEEANLIDHPDAFNDLLISDIKLNKEIVQAPVDFVIGNRSKTEIET
IYFSKVLVEGIIIVFENQELDLDIRIFVDTDADEILRMVRDVQEGSDVDCIMS
RYLSMWKPEKEIEPTKYADIIVHGNVRQNVVNIISQIKHLENALSDETYYM
VNSK"
complement(1047. .2021)
/gene="CP0012"
complement(1047. .2021)
/gene="CP0012"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA037908.1"
/db_xref="GI:7188951"
/translation="MEKKYLAAYVYITVDNHEEIALHKKFLEDVDSVIVISQ
GINDQKSHYEPHAEIKMQLKERNPSKIKFHIHKNFPIITVYKRELALGCE
VILSKQAKHISPOENMEKLEENCLLIDVNNYEMKIGHEFNATLPDIOTFERPEVA
EKLAQCDPTTPMVCCTGICELYSVPLEKKEVYQDGGVIAVQOQVGTGKM
LGKLFVDFDLAIPIDSDPVAIEACCHCQTFSDAYNCANTCNALFLCODECIH
QHGGCGEGCSFPRVAFDPSRGNKFFRZHLCEISENSESACCL1"
2238. .2867
/gene="CP0013"
2238. .2867
/gene="CP0013"
/note="similar to SP:P81288; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S4"
/protein_id="AA037909.1"
/db_xref="GI:7188952"
/translation="MAYCGPQVRRVARRFGANFGRSNPLKKPPPGQGHQKIKK
SGYQLQEEKQKACVGMKQVKAPEVHKQNVQVAFELFECELDNNVYRM
DFKTIQFAAQVLAHGHILVNGRRVDRSPFLRPGMQLSLKSKSLQSDALESKD
ESSIPSYISLDTGKFGELLAVSPQQLHAQLPLPINISVVCFLSHRT"
complement(2986. .3867)
/gene="CP0014"
complement(2986. .3867)
/gene="CP0014"
/note="similar to SP:P12638 GB:M22591 PID:146954
PID:405898 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="endonuclease IV"
/protein_id="AA037910.1"
/db_xref="GI:7188953"
/translation="MKVLPSPPLDLCATSTAGGLKNAIYSGRDIGASTVQITFANQ
RQWRALKEEVEDFKALKTDLTYSIMSHAGYLINPGAPDVILEKSRIGIYQIL
DCITLGSFYNFHFGALKSKEDCKNKIVSSFSQAPLDFSPFVLLVETTAGGT
TCLSGPEELGYLVQNLKNQIPIGVDTCHIFAAGYDITSPQGWEDVINEFDEYVGLS
YLRAFHLNDSMFPLGANKDRHAPLGEYIGKESFKELMTDTRTKPIKYLETFGPPN
WQKEIGELLKFSKRD"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:53:56 ; Search time 17.7616 Seconds

(without alignments)
3525.629 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MWNPIGPGIDETERTPAD.....SLFIQQLVNLGSLGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	2 D72042	conserved hypothet
2	3187	100.0	651	2 E96581	CHLPN 76 kda homol
3	2205	89.2	715	2 I40729	hypothetical 76K p
4	899.5	28.2	647	2 G71490	hypothetical prote
5	232.5	7.3	2055	2 T31110	extracellular matr
6	232	7.3	1822	2 S33441	EF protein - strep
7	223.5	7.0	2481	2 D90011	FmtB protein [impo
8	221.5	7.0	971	2 B30835	probable tail fibe
9	221.5	7.0	973	2 C85693	probable membrane
10	220	6.9	6713	2 B99921	hypothetical prote
11	210	6.6	1122	2 G4887	probable tail fibe
12	210	6.6	2271	2 F90073	hypothetical prote
13	208.5	6.5	4776	2 E95206	cell wall surface
14	199	6.2	1365	2 T30822	Imp1 protein - Myc
15	196.5	6.2	1829	2 T24583	hypothetical prote
16	196.5	6.2	2232	2 T34434	hypothetical prote
17	192	6.0	1156	2 T34852	probable secreted
18	190.5	6.0	1063	2 D86731	hypothetical prote
19	189	5.9	2155	2 AD2742	conserved hypothet
20	189	5.9	2155	2 C97523	hypothetical prote
21	187	5.9	1147	2 T35781	hypothetical prote
22	186.5	5.9	1306	2 S25370	hypothetical prote
23	183	5.7	2186	2 H89960	MSB2 protein - yea
24	182.5	5.7	641	2 C82206	hypothetical prote
25	182.5	5.7	1238	2 T03465	methy1-accepting c
26	182.5	5.7	2541	2 S11661	probable exonuclea
27	182	5.7	1269	2 F47330	tailin - mouse
28	178.5	5.6	1566	2 A43607	probable myosin he
29	177	5.6	2090	2 S26058	cell surface antig
					probable transform

30	176.5	5.5	545	2 E94327	Hcr7 transducer [i
31	176.5	5.5	545	2 T46811	halobacterial tran
32	176.5	5.5	3890	2 C89921	hypothetical prote
33	175.5	5.5	1104	2 S9310	probable membrane
34	175.5	5.5	5327	2 T13564	microtubule-associ
35	175	5.5	627	2 F94194	Hcr14 transducer [
36	175	5.5	891	2 G41662	130K surface exclu
37	175	5.5	1302	1 JC6009	surface-located me
38	173.5	5.4	1561	1 S06839	surface antigen sp
39	173	5.4	1072	2 A86827	hypothetical prote
40	173	5.4	2285	2 T12796	probable transglyc
41	172.5	5.4	1128	2 T30296	R27-2 protein - Tr
42	172.5	5.4	1601	2 AB1730	hypothetical prote
43	171.5	5.4	990	2 I51618	nucleolar phosphop
44	171	5.4	1787	2 AS1360	probable tape-meas
45	170.5	5.3	583	2 S67571	hypothetical prote

ALIGNMENTS

RESULT 1

D72042
conserved hypothetical protein CP0018 [imported] - Chlamydia pneumoniae (strains CW
N:Alternate names: chlpn 76 kda homolog 1 (cr622); hypothetical protein CPn0728
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C/Accession: D72042; D81623
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.
R:Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72042
A:Molecule type: DNA
A:Residues: 1-651 <ARN>
A:Cross-references: GB:AE001654; GB:AE001363; NID:94377031; PIDN:AA018867.1; PID:94377031;
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gyll, S.R.; Heidelberg, J.F.; White, O.; Hick
A: C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salz
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR3
A:Reference number: A81500; MUID:20150255; PMID:10684955
A:Accession: D81623
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-651 <REA>
A:Cross-references: GB:AE002165; GB:AE002161; NID:G7188948; PIDN:AAF37914.1; PID:G7188948;
A:Experimental source: strain AR39, HL cells
C:Comment: This sequence was originally identified as homologous to part of a sequence
PIR:H71490).
C:Genetics:
A:Gene: CPn0728; CP0018

Query Match 100.0%; Score 3187; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 1.9e-148;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MWNPIGPGIDETERTPPADLSAQGLESAANKSAQAQIAGAEAKPKESKTSVERWSI	60
Db	1	MWNPIGPGIDETERTPPADLSAQGLESAANKSAQAQIAGAEAKPKESKTSVERWSI	60
Qy	61	LRSAVNAWLSADKLGITFASNSSSSRSDVDSTTATATPPPTFDYKTAQAYDT	120
Db	61	LRSAVNAWLSADKLGITFASNSSSSRSDVDSTTATATPPPTFDYKTAQAYDT	120
Qy	121	IFTSTSLADIQAALVSLQDAVTNIKDTATDETAIAAEWETKNADAVKGAQITELAKY	180
Db	121	IFTSTSLADIQAALVSLQDAVTNIKDTATDETAIAAEWETKNADAVKGAQITELAKY	180
Qy	181	ASDNQAILDSGLKTSFDLLQAALLOSANNKKAELKEMQDNFVVGKTPAIAQSLVD	240
Db	181	ASDNQAILDSGLKTSFDLLQAALLOSANNKKAELKEMQDNFVVGKTPAIAQSLVD	240

241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISNIDSAKAAIATAKTQIAEAQ 300
 Db
 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISNIDSAKAAIATAKTQIAEAQ 300
 Qy
 301 KKFDPSPILQAEQWVIOAEKDLNKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMML 360
 Db
 301 KKFDPSPILQAEQWVIOAEKDLNKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMML 360
 Qy
 361 DDAENETASILMSGFRMIHMFNTENPDQAOELAAQAAKAAAGDSDSAAALADAQK 420
 Db
 361 DDAENETASILMSGFRMIHMFNTENPDQAOELAAQAAKAAAGDSDSAAALADAQK 420
 Qy
 421 ALEAALGKAGQOQGLNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSTGSDYKT 480
 Db
 421 ALEAALGKAGQOQGLNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSTGSDYKT 480
 Qy
 481 QISAGYDAYKSIINDAYGEARNDATRDVINNVSTPALTRSPVPRATEARGPEKTDQALRV 540
 Db
 481 QISAGYDAYKSIINDAYGEARNDATRDVINNVSTPALTRSPVPRATEARGPEKTDQALRV 540
 Qy
 541 ISGNRTLDGVYSQVVSALQSVMOIIOQNPOANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
 Db
 541 ISGNRTLDGVYSQVVSALQSVMOIIOQNPOANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
 Qy
 601 STQKFTAKLESFAEGSRTAAEIKALSFETNSLFIQOVLNIGSLYGYLQ 651
 Db
 601 STQKFTAKLESFAEGSRTAAEIKALSFETNSLFIQOVLNIGSLYGYLQ 651
 Qy

RESULT 2
 E86581
 CHLPN 76 kDa homolog_1 (C822) [imported] - Chlamydomophila pneumoniae (strain J138)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: E86581
 R:Shirai, M.; Hirakawa, H.; Kimoko, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
 Nucleic Acids Res 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20040349; PMID:10871362
 A:Accession: E86581
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-651 <S>
 A:Cross-references: GB:BA000008; NID:58979100; PIDN:BAA98935.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: CPJ0728

Query Match 100.0%; Score 3197; DB 2; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.9e-148;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVNPIGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
 Db 1 MVNPIGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
 Qy 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
 Db 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
 Qy 121 IFTSTSLADIQAALVSLQDAVTNKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 180
 Db 121 IFTSTSLADIQAALVSLQDAVTNKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 180
 Qy 181 ASDNQAILDSLGKLTSPDLLQALLOSANNVNNKAAELLEMQDNVPVPGKTPAIAQSLVD 240
 Db 181 ASDNQAILDSLGKLTSPDLLQALLOSANNVNNKAAELLEMQDNVPVPGKTPAIAQSLVD 240
 Qy 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISNIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISNIDSAKAAIATAKTQIAEAQ 300
 Qy 301 KKFDPSPILQAEQWVIOAEKDLNKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMML 360

301 KKFDPSPILQAEQWVIOAEKDLNKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMML 360
 Qy
 361 DDAENETASILMSGFRMIHMFNTENPDQAOELAAQAAKAAAGDSDSAAALADAQK 420
 Db
 361 DDAENETASILMSGFRMIHMFNTENPDQAOELAAQAAKAAAGDSDSAAALADAQK 420
 Qy
 421 ALEAALGKAGQOQGLNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSTGSDYKT 480
 Db
 421 ALEAALGKAGQOQGLNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSTGSDYKT 480
 Qy
 481 QISAGYDAYKSIINDAYGEARNDATRDVINNVSTPALTRSPVPRATEARGPEKTDQALRV 540
 Db
 481 QISAGYDAYKSIINDAYGEARNDATRDVINNVSTPALTRSPVPRATEARGPEKTDQALRV 540
 Qy
 541 ISGNRTLDGVYSQVVSALQSVMOIIOQNPOANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
 Db
 541 ISGNRTLDGVYSQVVSALQSVMOIIOQNPOANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
 Qy
 601 STQKFTAKLESFAEGSRTAAEIKALSFETNSLFIQOVLNIGSLYGYLQ 651
 Db
 601 STQKFTAKLESFAEGSRTAAEIKALSFETNSLFIQOVLNIGSLYGYLQ 651
 Qy

RESULT 3
 I40729
 hypothetical 76K protein - Chlamydomophila pneumoniae (strain AR39)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
 C:Accession: I40729
 R:Perez-Melgosa, M.; Kud, C.
 Infect. Immun. 62, 880-886, 1994
 A:Title: Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-kDa
 A:Reference number: I40729; MUID:94156481; PMID:7509320
 A:Accession: I40729
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-715 <RES>
 A:Cross-references: GB:I23921; NID:9435961; PIDN:AAA23117.1; PID:9435962
 A:Experimental source: strain AR-39
 C:Comment: This is the hypothetical translation of a sequence that was reported as two

Query Match 69.2%; Score 2205; DB 2; Length 715;
 Best Local Similarity 99.6%; Pred. No. 1.9e-100;
 Matches 452; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVNPIGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 60
 Db 257 LVNPIGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTSVERWSI 316
 Qy 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
 Db 317 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 376
 Qy 121 IFTSTSLADIQAALVSLQDAVTNKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 180
 Db 377 IFTSTSLADIQAALVSLQDAVTNKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 436
 Qy 181 ASDNQAILDSLGKLTSPDLLQALLOSANNVNNKAAELLEMQDNVPVPGKTPAIAQSLVD 240
 Db 437 ASDNQAILDSLGKLTSPDLLQALLOSANNVNNKAAELLEMQDNVPVPGKTPAIAQSLVD 49
 Qy 241 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISNIDSAKAAIATAKTQIAEAQ 30
 Db 497 QTDATATQIEKDGNAIRDAYFAGONASGAVENAKNSNISNIDSAKAAIATAKTQIAEAQ 55
 Qy 301 KKFDPSPILQAEQWVIOAEKDLNKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMML 360
 Db 557 KKFDPSPILQAEQWVIOAEKDLNKPADGSDVNPFGTTVGGSKQGGSSIGSIRVSMML 616
 Qy 361 DDAENETASILMSGFRMIHMFNTENPDQAOELAAQAAKAAAGDSDSAAALADAQK 420
 Db 617 DDAENETASILMSGFRMIHMFNTENPDQAOELAAQAAKAAAGDSDSAAALADAQK 676

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	2204	100.0	651	2	D72042		conserved hypotchet
2	2204	100.0	651	2	E86581		CHLPN 76 kDa homol
3	2201	99.9	715	2	I40729		hypothetical 76K p
4	461.5	20.9	647	2	G71490		hypothetical prote
5	197	8.9	971	2	B90835		probable tail fibe
6	197	8.9	973	2	B85693		probable membran
7	192.5	8.7	2055	2	T31110		extracellular matr
8	185.5	8.4	1822	2	S33441		EF protein - Strept
9	184	8.3	1122	2	G64887		probable tail fibe
10	181	8.2	2481	2	D90011		FmcB protein [mpo
11	176.5	8.0	545	2	E84327		Htr7 transducer [i
12	176.5	8.0	545	2	T46811		halobacterial tran
13	175	7.9	627	2	F84194		Htr14 transducer [
14	173	7.8	2271	2	F90073		hypothetical prote
15	172.5	7.8	1128	2	T30296		K27-2 protein - Tr
16	168.5	7.6	2232	2	T34434		hypothetical prote
17	169	7.6	993	2	C90072		hypothetical prote
18	164	7.4	6713	2	B89921		hypothetical prote
19	163.5	7.4	2155	2	AD2742		conserved hypothet
20	163.5	7.4	2155	2	C97523		hypothetical prote
21	160.5	7.3	544	2	T44938		transducer protein
22	160	7.3	536	1	A47190		transducer protein
23	160	7.3	536	2	E84318		Htr1 transducer [i
24	158.5	7.2	4776	2	E95206		cell wall surface
25	157	7.1	641	2	C83206		methyl-accepting c
26	156	7.1	810	2	T46810		halobacterial tran
27	156	7.1	810	2	F84327		Htrs transducer [i
28	156	7.1	2186	2	H89960		hypothetical prote
29	154.5	7.0	1238	2	T03465		probable exonuclea

QY	181	ASDNOAILDSLGKLTSTFDLLQALILQSVANNKKAELLKEMQDNPVPVGKTPAIAQSLVD	240
	181	ASDNOAILDSLGKLTSTFDLLQALILQSVANNKKAELLKEMQDNPVPVGKTPAIAQSLVD	240
Db	181	ASDNOAILDSLGKLTSTFDLLQALILQSVANNKKAELLKEMQDNPVPVGKTPAIAQSLVD	240

probable exonuclea

241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAIATAKTQIAEAQ 300
 |||||
 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAIATAKTQIAEAQ 300
 |||||
 301 KKFPPSPILQEAQWVIOAEKDLKNIKPADGSDVNPFGTTVGSKQGGSSIGSIRVSMML 360
 |||||
 301 KKFPPSPILQEAQWVIOAEKDLKNIKPADGSDVNPFGTTVGSKQGGSSIGSIRVSMML 360
 |||||
 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 |||||
 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 |||||
 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
 |||||
 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
 |||||

RESULT 2
 E86581
 CHLPN 76 kDa homolog 1 (CT622) [imported] - Chlamydomophila pneumoniae (strain J138)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: E86581
 R:Shirai, M.; Hirakawa, H.; Kinoy, M.; Tabuchi, M.; Kishi, F.; Cuchi, K.; Shiba, T.; Is
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-651 <SVC>
 A:Cross-references: GB:BA000008; NID:g8979100; PIDN:BAA98935.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: CPJ0728

Query Match 100.0%; Score 2204; DB 2; Length 651;
 Best Local Similarity 100.0%; Pred. No. 7.2e-110;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 60
 |||||
 Db 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 60
 |||||
 QY 61 LRSVNAALMSLADKLGIASSNSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 120
 |||||
 Db 61 LRSVNAALMSLADKLGIASSNSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 120
 |||||
 QY 121 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 180
 |||||
 Db 121 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 180
 |||||
 QY 181 ASDNQAILDSLGKLTSTFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 |||||
 Db 181 ASDNQAILDSLGKLTSTFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 |||||
 QY 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAIATAKTQIAEAQ 300
 |||||
 Db 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAIATAKTQIAEAQ 300
 |||||
 QY 301 KKFPPSPILQEAQWVIOAEKDLKNIKPADGSDVNPFGTTVGSKQGGSSIGSIRVSMML 360
 |||||
 Db 301 KKFPPSPILQEAQWVIOAEKDLKNIKPADGSDVNPFGTTVGSKQGGSSIGSIRVSMML 360
 |||||
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 |||||
 Db 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 |||||
 QY 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
 |||||
 Db 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
 |||||

RESULT 3
 I40729
 Hypothetical 76k protein - Chlamydomophila pneumoniae (strain AR39)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
 C:Accession: I40729
 R:Perez-Melgosa, M.; Kuo, C.
 Infect. Immun. 62, 880-886, 1994
 A:Title: Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-kil
 A:Reference number: I40729; MUID:94156481; PMID:7509320
 A:Accession: I40729
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-715 <RES>
 A:Cross-references: GB:I23921; NID:g435961; PIDN:AAK23117.1; PID:g435962
 A:Experimental source: strain AR-39
 C:Comment: This is the hypothetical translation of a sequence that was reported as two s

Query Match 99.9%; Score 2201; DB 2; Length 715;
 Best Local Similarity 99.8%; Pred. No. 1.2e-109;
 Matches 451; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 60
 |||||
 Db 257 LVPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTSVERWSI 316
 |||||
 QY 61 LRSVNAALMSLADKLGIASSNSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 120
 |||||
 Db 317 LRSVNAALMSLADKLGIASSNSSTSRSDVSTTATPTPPPTFDDYKTOAQATYDT 376
 |||||
 QY 121 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 180
 |||||
 Db 377 IFTSTSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKGAQITELAKY 436
 |||||
 QY 181 ASDNQAILDSLGKLTSTFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 |||||
 Db 437 ASDNQAILDSLGKLTSTFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 496
 |||||
 QY 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAIATAKTQIAEAQ 300
 |||||
 Db 497 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAIATAKTQIAEAQ 556
 |||||
 QY 301 KKFPPSPILQEAQWVIOAEKDLKNIKPADGSDVNPFGTTVGSKQGGSSIGSIRVSMML 360
 |||||
 Db 557 KKFPPSPILQEAQWVIOAEKDLKNIKPADGSDVNPFGTTVGSKQGGSSIGSIRVSMML 616
 |||||
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 |||||
 Db 617 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAAGDDSAALADAQK 676
 |||||
 QY 421 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 452
 |||||
 Db 677 ALEAALGKAGQCGGILNALGQIASAAVVSAGV 708
 |||||

RESULT 4
 G71490
 Hypothetical protein CT622 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
 N:Alternate names: chlpn 76kDa homolog CT622
 C:Species: Chlamydia trachomatis
 C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
 C:Accession: G71490
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell
 Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tr
 A:Reference number: A71570; MUID:99000809; PMID:9784136
 A:Accession: G71490
 A:Molecule type: DNA
 A:Residues: 1-647 <ARN>
 A:Cross-references: GB:AE001333; GB:AE001273; NID:g3329068; PIDN:AAK68226.1; PID:g3329
 A:Experimental source: serotype D, strain UW-3/Cx
 C:Comment: This sequence was originally identified as homologous to part of a sequence
 FIR:E72042).

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:04 ; Search time 560.656 Seconds
(without alignments)
11032.393 Million cell updates/sec

Title: US-10-608-559-5
Perfect score: 1456
Sequence: 1 ataaatctttaaaacagg.....ctgttgagcgcaggagta 1456

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04.*

- 1: Geneseq_1980s.*
- 2: Geneseq_1990s.*
- 3: Geneseq_2000s.*
- 4: Geneseq_2001as.*
- 5: Geneseq_2001bs.*
- 6: Geneseq_2002s.*
- 7: Geneseq_2003as.*
- 8: Geneseq_2003bs.*
- 9: Geneseq_2003cs.*
- 10: Geneseq_2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1456	100.0	1456	3	AAD02065
2	1456	100.0	2238	3	AAD02066
3	1455	99.9	2156	3	AAD02063
4	1444	99.2	110000	2	AAX91990_08
5	1355	93.1	1956	4	AAS57031 C. pneumo
6	1355	93.1	1956	6	ABL92612 Chlamydia
7	1355	93.1	1956	6	ABL91190 Chlamydia
8	1355	93.1	1956	9	ADD42825 Chlamydia
9	1151	79.1	1852	3	AAD02064 5'-trunca
10	170	11.7	1550	3	AAX28411 Chlamydia
11	148.4	10.2	150	2	AAV16207 Part of t
12	55.4	3.8	2000	7	ADA71938 Rice gene
13	54.4	3.7	1941	6	ABL92619 Chlamydia
14	54.4	3.7	1944	9	ADD43867 Chlamydia
15	51.8	3.6	110000	2	AAX201425_07
16	51.8	3.6	110000	2	AAZ01425_08
17	51.6	3.5	2205	8	ADA28999 DNA encod
18	51.4	3.5	1537	3	AAG63292 C. tracho
19	51.4	3.5	1537	4	AAS56196 Chlamydia
20	51.4	3.5	1537	6	ABL92425 Chlamydia
21	49	3.4	896	2	AAQ68903 PspA prot
22	49	3.4	946	2	AAX33124 Streptoco
23	49	3.4	957	2	AAT61726 Streptoco

24	49	3.4	1860	7	ACA49648
25	49	3.4	1860	2	AA161725
26	49	3.4	1990	7	ABX95373
27	49	3.4	2085	2	AAQ78131
28	49	3.4	2085	2	AAT08979
29	49	3.4	2085	2	AAT07178
30	49	3.4	2085	2	AAV33264
31	49	3.4	2085	2	AAV39470
32	49	3.4	2085	2	AAZ25063
33	49	3.4	2085	2	AAV84069
34	49	3.4	2085	2	AAQ02012
35	48.6	3.3	1171	4	AAS56981
36	48.6	3.3	1171	9	ADD42775
37	48.6	3.3	1834	4	AAS57006
38	48.6	3.3	1834	9	ADD42800
39	48.6	3.3	1983	4	AAS56996
40	48.6	3.3	1983	9	ADD42790
41	47.4	3.3	956	7	ABX95374
42	47.4	3.3	2085	2	AAQ28674
43	47.2	3.2	110000	6	ABA90521_07
44	47	3.2	9373	6	ABL33291
45	47	3.2	9373	6	ABK31317

ALIGNMENTS

RESULT 1
AAD02065
ID AAD02065 standard; DNA; 1456 BP.

XX AC AAD02065;

DT 26-MAR-2001 (first entry)

DE 3'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.

XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; truncation mutant; ds.

OS Chlamydia pneumoniae.

XX Synthetic.

XX Key Location/Qualifiers
FT CDS 101..1456

FT /*tag= a

FT /product= "3'-truncated Chlamydia pneumoniae 76KDa

FT /protein=

FT /note= "The coding region does not include stop codon"

FT /partial

XX WO2000066739-A2.

XX 09-NOV-2000.

XX 03-MAY-2000; 2000WO-CA000511.

XX 03-MAY-1999; 99US-0132270P.

XX 30-JUN-1999; 99US-0141276P.

XX (AVET) AVENTIS PASTEUR LTD.

XX WPI; 2000-687542/67.

XX P-PSDB; AAY71956.

XX Murdin AD, Oomen RP, Wang J, Dunn P;
P-Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.

XX Claim 2c; Page 102-104; 112pp; English.

Handwritten signature

XX The present sequence is a DNA coding for 3'-truncated Chlamydia
CC pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the
CC diagnosis, prevention and treatment of C. pneumoniae infections (e.g.
CC pneumonia, upper respiratory tract disease, bronchitis, sinusitis and
CC acute respiratory disease such as cough, sore throat, hoarseness, fever;
CC and abnormal chest sounds on auscultation). C. pneumoniae sequence is
CC also used as vaccines for immunising humans against diseases caused by C.
CC pneumoniae
XX

SQ Sequence 1456 BP; 452 A; 308 C; 331 G; 365 T; 0 U; 0 Other;
Query Match 100.0%; Score 1456; DB 3; Length 1456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 60
DB 1 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 60
QY 61 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 120
DB 61 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 120
QY 121 AGGTCCTATAGACGAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 180
DB 121 AGGTCCTATAGACGAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTATATTTT 180
QY 181 GCGAGTCAGCAATTAAGAGTCGCAAGCTCAAGAGTACAGAGTACAGAGTACAG 240
DB 181 GCGAGTCAGCAATTAAGAGTCGCAAGCTCAAGAGTACAGAGTACAGAGTACAG 240
QY 241 TAAAGATCTAAGACCGATTTCTAGAGGATGAGGATCTTTCGCTTCTGAGTGAATC 300
DB 241 TAAAGATCTAAGACCGATTTCTAGAGGATGAGGATCTTTCGCTTCTGAGTGAATC 300
QY 301 TCTCATGAGTCGGCAGATAAGCTGGGATTTGCTTCTAGTAAGCTGCTTCTACTAG 360
DB 301 TCTCATGAGTCGGCAGATAAGCTGGGATTTGCTTCTAGTAAGCTGCTTCTACTAG 360
QY 361 CAGATCTGACGAGCTGAGCTCAAGCAGCAGCGACCTTACGCTTCTCCACCCAGTT 420
DB 361 CAGATCTGACGAGCTGAGCTCAAGCAGCAGCGACCTTACGCTTCTCCACCCAGTT 420
QY 421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTTCTTACCTCAACATCACT 480
DB 421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTTCTTACCTCAACATCACT 480
QY 481 AGCTGACATACAGCTCTTGGTGGAGCTTCCAGGATGCTGCTCACTAATATAAGGATAC 540
DB 481 AGCTGACATACAGCTCTTGGTGGAGCTTCCAGGATGCTGCTCACTAATATAAGGATAC 540
QY 541 AGCGGCTACTGATGAGAAACCGCAATCGCTGGCGAGTGGAACTAAGATGCCGATGC 600
DB 541 AGCGGCTACTGATGAGAAACCGCAATCGCTGGCGAGTGGAACTAAGATGCCGATGC 600
QY 601 AGTTAAAGTTGGCGCGCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 660
DB 601 AGTTAAAGTTGGCGCGCAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 660
QY 661 TCTTGACTCTTTAGTAACTGACTTCTTCCGACTCTTACAGGCTGCTTCTTCCCAATC 720
DB 661 TCTTGACTCTTTAGTAACTGACTTCTTCCGACTCTTACAGGCTGCTTCTTCCCAATC 720
QY 721 TGTAGCAACATTAAGAGAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGT 780
DB 721 TGTAGCAACATTAAGAGAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGT 780
QY 781 CCCAGGAAACCGCTGCAATTTGCTCAATTTTGTAGTATGATCAGAGATGTAAGCGAC 840
DB 781 CCCAGGAAACCGCTGCAATTTGCTCAATTTTGTAGTATGATCAGAGATGTAAGCGAC 840
QY 841 ACAGATAGAAAGATGGAATGGCATAGGATGATATTTTGCAGGACAGACGCTAG 900

DB 841 ACAGATAGAAAGATGGAATGGCATAGGATGATATTTTGCAGACAGACGCTAG 900
QY 901 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTGAGTAAAGC 960
DB 901 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTGAGTAAAGC 960
QY 961 AGCAATCGTCTACTGCTAAGACACAAATAGCTTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
DB 961 AGCAATCGTCTACTGCTAAGACACAAATAGCTTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
QY 1021 AATTTCTTCAAGAGCGAAACAAATGGTAATACAGGCTCAGAAAAGTCTTAAAAATATCAA 1080
DB 1021 AATTTCTTCAAGAGCGAAACAAATGGTAATACAGGCTCAGAAAAGTCTTAAAAATATCAA 1080
QY 1081 ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGACACA 1140
DB 1081 ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGACACA 1140
QY 1141 AGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
DB 1141 AGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
QY 1201 CGCTTCCATTTTGTATGCTGGGTTTCGTCAGATGATTCACATGTTCAATAGGAAAATCC 1260
DB 1201 CGCTTCCATTTTGTATGCTGGGTTTCGTCAGATGATTCACATGTTCAATAGGAAAATCC 1260
QY 1261 TGATTTCTCAAGTGTCCCAACAGGAGCTCGACACAGCTAGACAGCGAAAGCCGCTGG 1320
DB 1261 TGATTTCTCAAGTGTCCCAACAGGAGCTCGACACAGCTAGACAGCGAAAGCCGCTGG 1320
QY 1321 AGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGGCTTAGG 1380
DB 1321 AGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGGCTTAGG 1380
QY 1381 TAAAGCTGGCAACCAACAGGCGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGT 1440
DB 1381 TAAAGCTGGCAACCAACAGGCGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGT 1440
QY 1441 TGTGAGCGCAGAGTA 1456
DB 1441 TGTGAGCGCAGAGTA 1456

RESULT 2
AAD02066
ID AAD02066 standard; DNA; 2238 BP.
XX
AC AAD02066;
XX 26-MAR-2001 (first entry)
XX
C. pneumoniae 76 kDa protein truncation mutant fusion gene.
76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
upper respiratory tract disease; bronchitis; sinusitis;
acute respiratory disease; cough; sore throat; hoarseness; fever;
vaccine; immunisation; treatment; truncation mutant; fusion gene; ds.
Chlamydia pneumoniae.
Synthetic.
Key Location/Qualifiers
misc_feature 1..665
/tag= a
/note= "This part of the sequence is unrelated to C.
pneumoniae 76 kDa gene"
CDS 766..2238
/tag= b
/product= "Truncated Chlamydia pneumoniae 76KDa protein"
misc_feature 2122..2238
/tag= c
/note= "This part of the sequence is unrelated to C."

pneumoniae 76 kDa gene"

FT XX WO200066739-A2.
PN XX 09-NOV-2000.
XX XX 03-MAY-2000; 2000WO-CA000511.
XX XX 03-MAY-1999; 99US-0132270P.
PR 30-JUN-1999; 99US-0141276P.
XX XX (AVET) AVENTIS PASTEUR LTD.
PA XX
XX XX Murdin AD, Oomen RP, Wang J, Dunn P;
PI XX WPI; 2000-687542/67.
XX XX P-PSDB; ANY71957.
DR XX
XX XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
PT for vaccinating against Chlamydia infections.
XX XX
XX XX Claim 32; Fig 3; 112pp; English.
PS XX
XX XX The present sequence is a DNA coding for a fusion protein comprising a
CC truncated Chlamydia pneumoniae 76 kDa protein and vector-encoded
CC residues. C. pneumoniae 76 kDa protein is used in the diagnosis,
CC prevention and treatment of C. pneumoniae infections (e.g. pneumonia,
CC upper respiratory tract disease, bronchitis, sinusitis and acute
CC respiratory disease such as cough, sore throat, hoarseness, fever; and
CC abnormal chest sounds on auscultation). C. pneumoniae sequence is also
CC used as vaccines for immunising humans against diseases caused by C.
XX pneumoniae
SQ Sequence 2238 BP; 689 A; 459 C; 497 G; 593 T; 0 U; 0 Other;
Query Match 100.0%; Score 1456; DB 3; Length 2238;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTACAGAGCTTTTATTTTATTTT 60
DB 666 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTACAGAGCTTTTATTTTATTTT 725
QY 61 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTACAGAGCTTTTATTTTATTTT 120
DB 726 ATAAATCTTTAAACACAGGCTCGCATTAATATTAGTACAGAGCTTTTATTTTATTTT 785
QY 121 AGTCTATAGACGAAACAGACGACACCTCCCGCAGATCTTCTGCTCAAGGATTGGA 180
DB 786 AGTCTATAGACGAAACAGACGACACCTCCCGCAGATCTTCTGCTCAAGGATTGGA 845
QY 181 GCGAGTGCAGCAATAAGAGTGCAGGAGCTCAAGAAATAGAGTGCAGGAGCTTAAGCC 240
DB 846 GCGAGTGCAGCAATAAGAGTGCAGGAGCTCAAGAAATAGAGTGCAGGAGCTTAAGCC 905
QY 241 TAAAGATCTAAGACGATCTGTAGAGGATGGAGCATCTTGGCTTGTGAGTGAATGC 300
DB 906 TAAAGATCTAAGACGATCTGTAGAGGATGGAGCATCTTGGCTTGTGAGTGAATGC 965
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAAACAGCTGCTTCTACTAG 360
DB 966 TCTCATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAAACAGCTGCTTCTACTAG 1025
QY 361 CAGATCTGCAGACGCTGGAATCAACGACAGCGACCGCAGCTAGCCCTCTCCACCCACGTT 420
DB 1026 CAGATCTGCAGACGCTGGAATCAACGACAGCGACCGCAGCTAGCCCTCTCCACCCACGTT 1085
QY 421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT 480
DB 1086 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT 1145
QY 481 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACCTAATATAAGGATAC 540

DB 1146 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGTCTACTAATATAAGGATAC 1205
QY 541 AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACCTAAGAAATGCCGATGC 600
DB 1206 AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACCTAAGAAATGCCGATGC 1265
QY 601 AGTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGAT 660
DB 1266 AGTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGAT 1325
QY 661 TCTTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
DB 1326 TCTTGACTCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 1385
QY 721 TGTAGCAAAACAATAAAGAGAGTGCAGTCTTTAAAGAGATCGAAGATACCCAGTAGT 780
DB 1386 TGTAGCAAAACAATAAAGAGAGTGCAGTCTTTAAAGAGATCGAAGATACCCAGTAGT 1445
QY 781 CCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
DB 1446 CCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 1505
QY 841 ACAGATAGAAAGATGGAATCGATTAGGATGCAATTTTTCAGGACAGAAACGCTAG 900
DB 1506 ACAGATAGAAAGATGGAATCGATTAGGATGCAATTTTTCAGGACAGAAACGCTAG 1565
QY 901 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTTCAGTAAAGC 960
DB 1566 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTTCAGTAAAGC 1625
QY 961 AGCAATCGTACTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCC 1020
DB 1626 AGCAATCGTACTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCC 1685
QY 1021 AATTCTTTCAAGACGGAACAAATGGTAATACAGGCTGAGAAAAGTCTTAAAAATATCAA 1080
DB 1686 AATTCTTTCAAGACGGAACAAATGGTAATACAGGCTGAGAAAAGTCTTAAAAATATCAA 1745
QY 1081 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGGCAACA 1140
DB 1746 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGGCAACA 1805
QY 1141 AGGAAGTAGTATTGCTGATGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
DB 1806 AGGAAGTAGTATTGCTGATGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1865
QY 1201 CGCTTCCATTTTGTGTTGCTGGTTTCGTCAGATGATTCATGTTCAATACGAAAATCC 1260
DB 1866 CGCTTCCATTTTGTGTTGCTGGTTTCGTCAGATGATTCATGTTCAATACGAAAATCC 1925
QY 1261 TGATTTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGAGCAGGAAAGCGGCTGG 1320
DB 1926 TGATTTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGAGCAGGAAAGCGGCTGG 1985
QY 1321 AGATGACAGTCTGCTGCGAGGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGG 1380
DB 1986 AGATGACAGTCTGCTGCGAGGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGG 2045
QY 1381 TAAAGCTGGGCAACAAAGGCGATACCTCAATGCTTTTAGGACAGATCGGTTCTGCTGCTGT 1440
DB 2046 TAAAGCTGGGCAACAAAGGCGATACCTCAATGCTTTTAGGACAGATCGGTTCTGCTGCTGT 2105
QY 1441 TGTAGCGCAGAGTA 1456
DB 2106 TGTAGCGCAGAGTA 2121

RESULT 3
AAD02063
ID AAD02063 standard; DNA; 2156 BP.
XX AC
XX AAD02063;

DT	15-SEP-2003 (revised)
DT	26-MAR-2001 (first entry)
XX	Chlamydia pneumoniae 76 kDa full-length protein encoding DNA.
XX	
XX	76 kDa protein; bactericidal; diagnosis; prevention; treatment;
KW	pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
KW	acute respiratory disease; cough; sore throat; hoarseness; fever;
KW	vaccine; immunisation; ds.
XX	Chlamydophila pneumoniae.
OS	
PH	Key Location/Qualifiers
FT	101..2056
CDS	/tag= a
FT	/product= "Chlamydia pneumoniae 76kDa protein"
XX	WO200065739-A2.
PX	
PN	09-NOV-2000.
PD	
XX	03-MAY-2000; 2000WC-CA000511.
PF	
XX	03-MAY-1999; 99US-0132270P.
PR	
XX	30-JUN-1999; 99US-0141276P.
PA	(AVET) AVENTIS PASTEUR LTD.
PI	Murdin AD, Oomen RP, Wang J, Dunn P;
PI	
XX	WPI; 2000-687542/67.
DR	P-PSDB; AAY71954.
XX	Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
PT	for vaccinating against Chlamydia infections.
XX	Claim 2a; Fig 1; 112pp; English.
XX	The present sequence is a DNA coding for Chlamydia pneumoniae 76 kDa
CC	protein. C. pneumoniae 76 kDa protein is used in the diagnosis,
CC	prevention and treatment of C. pneumoniae infections (e.g. pneumonia,
CC	upper respiratory tract disease, bronchitis, sinusitis and acute
CC	respiratory disease such as cough, sore throat, hoarseness, fever; and
CC	abnormal chest sounds on auscultation). C. pneumoniae sequence is also
CC	used as vaccines for immunising humans against diseases caused by C.
CC	pneumoniae. (Updated on 15-SEP-2003 to standardise OS field)
XX	
SQ	Sequence 2156 BP; 572 A; 461 C; 471 G; 552 T; 0 U; 0 Other;
	Query Match 99.9%; Score 1455; DB 3; Length 2156;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATAAATCTTTTAAAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT 60
DB	1 ATAAAACTTTTAAAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT 60
QY	61 ATAATAAATCTTAAAAGATTTTTTATTTTTCAGTTTTTATGGTTAAATCCTATTGGTCC 120
DB	61 ATAATAAATCTTAAAAGATTTTTTATTTTTCAGTTTTTATGGTTAAATCCTATTGGTCC 120
QY	121 AGGTCTTAGACGAACAAGACGCACACTCCCGCAGATCTTCTGCTCAGAGATTGGA 180
DB	121 AGGTCTTAGACGAACAAGACGCACACTCCCGCAGATCTTCTGCTCAGAGATTGGA 180
QY	181 GCGAGTGCCAGCAATTAAGAGTGGGAGCTCAAAGAATAGCAGGTGCGGAAGCTAAGCC 240
DB	181 GCGAGTGCCAGCAATTAAGAGTGGGAGCTCAAAGAATAGCAGGTGCGGAAGCTAAGCC 240
QY	241 TAAAGAAATCTAAGACCGGATCTCTAGACGAGATGAGCACTTTCGGTTCTGAGTGATGC 300
DB	241 TAAAGAAATCTAAGACCGGATCTCTAGACGAGATGAGCACTTTCGGTTCTGAGTGATGC 300

Db	1381	TAAAGCTGGGCAACAACAGGGCACTCAATGCTTTAGGACAGATCGCTCTCTGCTGCT	1440
Qy	1441	TGTGAGCGCAGGAGT	1455
Db	1441	TGTGAGCGCAGGAGT	1455
RESULT 4			
AAx91990_08/c			
Continuation (9 of 13) of AAx91990 from base 800001 (Nucleotide sequence of the complete			
WP Sequence split into 13 fragments LOCUS AAx91990 Accession AAx91990			
WP	Fragment Name	Begin	End
WP	AAx91990_00	1	110000
WP	AAx91990_01	100001	210000
WP	AAx91990_02	200001	310000
WP	AAx91990_03	300001	410000
WP	AAx91990_04	400001	510000
WP	AAx91990_05	500001	610000
WP	AAx91990_06	600001	710000
WP	AAx91990_07	700001	810000
WP	AAx91990_08	800001	910000
WP	AAx91990_09	900001	1010000
WP	AAx91990_10	1000001	1110000
WP	AAx91990_11	1100001	1210000
WP	AAx91990_12	1200001	12300025
Query Match 99.2%; Score 1444; DB 2; Length 110000;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
Qy	1	ATAAATCTTTAAACACGGCTCGCATTAATNTAGTGAGAGCTTTTTTTTATTTTTT	60
Db	28698	ATAAATCTTTAAACACGGCTCGCATTAATNTAGTGAGAGCTTTTTTTTATTTTTT	28639
Qy	61	ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	120
Db	28638	ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	28579
Qy	121	AGGTCCTATAGACGAAACAGACGACACCTCCCGCAGATCTTTCTGTCAGAGATTGGA	180
Db	28578	AGGTCCTATAGACGAAACAGACGACACCTCCCGCAGATCTTTCTGTCAGAGATTGGA	28519
Qy	181	GGCGAGTCAGCAATATAGAGTCGGAGAGCTCAAGAAATGACAGTGGGAGCTAAGCC	240
Db	28518	GGCGAGTCAGCAATATAGAGTCGGAGAGCTCAAGAAATGACAGTGGGAGCTAAGCC	28459
Qy	241	TAAAGAACTTAAGACCGATTCTGTAGACGATGGAGCATCTTGCCTTCTGCAAGTGAATGC	300
Db	28458	TAAAGAACTTAAGACCGATTCTGTAGACGATGGAGCATCTTGCCTTCTGCAAGTGAATGC	28399
Qy	301	TCTCATGAGTCGGCAGATAAGCTGGGATTTGCTTTCTAGTAACAGCTGCTCTTCTACTAG	360
Db	28398	TCTCATGAGTCGGCAGATAAGCTGGGATTTGCTTTCTAGTAACAGCTGCTCTTCTACTAG	28339
Qy	361	CAGATCTGACAGCTGGATCTCAACACAGCGACCGACCTAGCCCTCTCCACCCACAGTT	420
Db	28338	CAGATCTGACAGCTGGATCTCAACACAGCGACCGACCTAGCCCTCTCCACCCACAGTT	28279
Qy	421	TGATGATTATAGATCTCAAGCGCAACAGCTTACGATATCTTTTACTCTCAACATCACT	480
Db	28278	TGATGATTATAGATCTCAAGCGCAACAGCTTACGATATCTTTTACTCTCAACATCACT	28219
Qy	481	AGCTGACATACAGGCTGCTTTGGTGAGCCTCAGGATGCTGCTCACTAATATAAGGATAC	540
Db	28218	AGCTGACATACAGGCTGCTTTGGTGAGCCTCAGGATGCTGCTCACTAATATAAGGATAC	28159
Qy	541	AGCGGCTACTGATAGGAACCGCAATCGCTCGGAGTGGAACTTAAGATTCGCGATGC	600
Db	28158	AGCGGCTACTGATAGGAACCGCAATCGCTCGGAGTGGAACTTAAGATTCGCGATGC	28099
Qy	601	AGTTAAAGTTGGCGCAAAATTACGAATTAGCGAAATATGCTTCGGATAACCAAGCGAT	660

RESULT 5

AAS57031

ID AAS57031 standard; DNA; 1956 BP.

XX

AC AAS57031;

XX

DT 11-SEP-2003 (revised)

DT 16-JAN-2002 (first entry)

XX

C. pneumoniae DNA encoding the Ctr622 homologue CPn0728.

XX Chlamydia; ds; sexually transmitted disease; PID; antibacterial;

KW pelvic inflammatory disease; antigen; trachoma; gynecological;

KW acute respiratory tract infection; atherosclerosis; male infertility;

DT 29-1
DT 05-1

XX
DT 29-AUG-2003 (revised)
DT 05-JUN-2002 (first entry)

DT	29-AUG-2003	(REVISED)
DT	05-JUN-2002	(first entry)

DT	29-AUG-2003	(Revised)
DT	05-JUN-2002	(first entry)

XX Chlamydia pneumoniae DNA sequence SEQ ID NO:385.
 DE Chlamydia pneumoniae; Chlamydia; vaccine; detection; diagnosis; antigen;
 XX Chlamydia pneumoniae; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW Chlamydia pneumoniae; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW Chlamydia pneumoniae; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW Chlamydia pneumoniae; Chlamydia; vaccine; detection; diagnosis; antigen;
 XX Chlamydia pneumoniae; Chlamydia; vaccine; detection; diagnosis; antigen;
 OS Chlamydia pneumoniae.
 XX Chlamydia pneumoniae.
 PN WO200208267-A2.
 XX 31-JAN-2002.
 PD 20-JUL-2001; 2001WO-US023121.
 XX 20-JUL-2000; 2000US-00620417.
 XX 23-APR-2001; 2001US-00841122.
 PR (CORI-) CORIXA CORP.
 XX
 PA Filing SP, Skeiky YAW, Probst P, Bhatia A;
 XX WPI; 2002-179901/23.
 XX Novel compositions comprising Chlamydia Cap1 protein and its use in the
 PT treatment of Chlamydia infection.
 PT
 XX Disclosure; Page 336-337; 537pp; English.
 XX The present invention describes compositions comprising a Chlamydia Cap1
 CC protein and methods for the diagnosis and therapy of Chlamydia infection.
 CC Chlamydia DNA and protein sequences from the present invention can have
 CC antibacterial and immunostimulant activities, and can be used in
 CC vaccines. Compounds from the present invention can be used for eliciting
 CC an immune response, specifically stimulating a Chlamydia-specific T-cell
 CC response or inhibiting the development of a Chlamydia infection in an
 CC animal. Methods from the present invention can be used: for detecting the
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia
 CC infection. ABL92394 to ABL92709 and ABL94096 to ABL94374 represent
 CC sequences used in the exemplification of the present invention. (Updated
 CC on 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;
 Query Match 93.1%; Score 1355; DB 6; Length 1956;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 101 ATGGTTAATCTATGGTCCAGGTCCTATAGACGAAACAGAACACACCTCCGCGAGAT 160
 1 ATGGTTAATCTATGGTCCAGGTCCTATAGACGAAACAGAACACACCTCCGCGAGAT 60
 161 CTTTCTCTCAAGGATTGGAGCGAGTGCAGCAATTAAGAGTGGGAGCTCAAGAATA 220
 61 CTTTCTCTCAAGGATTGGAGCGAGTGCAGCAATTAAGAGTGGGAGCTCAAGAATA 120
 221 GCAGGTGGGAGCTTAAGCTTAAGATCTTAAGACCGATTCTAGAGCGATGAGGATC 280
 121 GCAGGTGGGAGCTTAAGCTTAAGATCTTAAGACCGATTCTAGAGCGATGAGGATC 180
 281 TTGGTCTTCAGTGAATGCTCATAGTCTGGCAGATAGCTGGGATTGCTTCTAGT 340
 181 TTGGTCTTCAGTGAATGCTCATAGTCTGGCAGATAGCTGGGATTGCTTCTAGT 240
 341 AACAGCTCGTCTTCTACTAGCAGATCTCAGACGCTGAGCTCAACGACGACGACGACCT 400
 241 AACAGCTCGTCTTCTACTAGCAGATCTCAGACGCTGAGCTCAACGACGACGACGACCT 300
 401 AGCGCTCTCCACCCACGTTGATGATTATAGATCAAGCTCAAGCGCAACAGCTTACGATCT 460
 301 AGCGCTCTCCACCCACGTTGATGATTATAGATCAAGCTCAAGCGCAACAGCTTACGATCT 360

QY 461 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
 DB 361 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
 QY 521 GTCACTAAATATAAGGATACAGCGGCTACTGTAGAGGAAACCGCAATCGCTCGGAGTGG 580
 DB 421 GTCACTAAATATAAGGATACAGCGGCTACTGTAGAGGAAACCGCAATCGCTCGGAGTGG 480
 QY 581 GAAACTAAGATGCCGATGCGATTAAAGTTGGCGGCGCAATTAACAGATTAAGCAATAT 640
 DB 481 GAAACTAAGATGCCGATGCGATTAAAGTTGGCGGCGCAATTAACAGATTAAGCAATAT 540
 QY 641 GCTTCGGATAACCAAGGATCTTGTACTCTTTAGGTAAACTGACTCTTCCTCGACCTCTTA 700
 DB 541 GCTTCGGATAACCAAGGATCTTGTACTCTTTAGGTAAACTGACTCTTCCTCGACCTCTTA 600
 QY 701 CAGGCTGCTCTTCTCCAACTCTGTAGCAACAATAACAAAGCAGCTGAGCTTTCTTAAGAG 760
 DB 601 CAGGCTGCTCTTCTCCAACTCTGTAGCAACAATAACAAAGCAGCTGAGCTTTCTTAAGAG 660
 QY 761 ATGCAAGATAACCCAGTAGTCCAGGAAACCGCTGCAATTCGCTCAATCTTTAGTTGAT 820
 DB 661 ATGCAAGATAACCCAGTAGTCCAGGAAACCGCTGCAATTCGCTCAATCTTTAGTTGAT 720
 QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGATGCATAT 880
 DB 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGATGCATAT 780
 QY 881 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACGATTAAGC 940
 DB 781 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACGATTAAGC 840
 QY 941 AACATAGATTACGCTAAAGCAGCAATCGTACTCTTAAGACACAAATAGCTGAGCTCAG 1000
 DB 841 AACATAGATTACGCTAAAGCAGCAATCGTACTCTTAAGACACAAATAGCTGAGCTCAG 900
 QY 1001 AAAAAGTTCGCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060
 DB 901 AAAAAGTTCGCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 960
 QY 1061 AAAGATCTTAAAAAATATCAAACTCGCAGATGGTTCGATTTCCAAATCCAGGAACTACA 1120
 DB 961 AAAGATCTTAAAAAATATCAAACTCGCAGATGGTTCGATTTCCAAATCCAGGAACTACA 1020
 QY 1121 GTTGGAGGCTCCAGACACAGGAGTAGTATTTGGTAGTATTCGTTTCCATGCTGTTA 1180
 DB 1021 GTTGGAGGCTCCAGACACAGGAGTAGTATTTGGTAGTATTCGTTTCCATGCTGTTA 1080
 QY 1181 GATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGTTCTGGGTTTCGTCAGATGATTCAC 1240
 DB 1081 GATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGTTCTGGGTTTCGTCAGATGATTCAC 1140
 QY 1241 ATGTTCAATACGGAATAATCTGATTTCAAGCTGCCCCAACAGGAGCTCGGACGACAGCT 1300
 DB 1141 ATGTTCAATACGGAATAATCTGATTTCTCAAGCTGCCCCAACAGGAGCTCGGACGACAGCT 1200
 QY 1301 AGAGCAGGAAAGCGCTGGAGATGACAGTGTCTGACGCTGCGGAGATGCTCAGAAA 1360
 DB 1201 AGAGCAGGAAAGCGCTGGAGATGACAGTGTCTGACGCTGCGGAGATGCTCAGAAA 1260
 QY 1361 GCTTTAGAAGCGGCTCTAGTAAAGCTGGGCAACACAGGGCATACTCAATGCTTTAGGA 1420
 DB 1261 GCTTTAGAAGCGGCTCTAGTAAAGCTGGGCAACACAGGGCATACTCAATGCTTTAGGA 1320
 QY 1421 CAGATCCCTTCTGCTGCTGTTGTAGCGCGAGGAGT 1455
 DB 1321 CAGATCCCTTCTGCTGCTGTTGTAGCGCGAGGAGT 1355

RESULT 7
 ABL91190
 ID ABL91190 standard; DNA; 1956 BP.
 XX

ABL91190; AC
XX XX

29-AUG-2003 (revised) DT
29-JUL-2002 (first entry) DD

Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.

Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
human respiratory disease; cardiovascular disease; atherosclerosis;
coronary artery disease; carotid artery stenosis; myocardial infarction;
cerebrovascular disease; aortic aneurysm; claudication; stroke;
strain CWL029; open reading frame; ORF; gene; ds.
XX KW
XX DE

Chlamydoxiphila pneumoniae.
OS

Key Location/Qualifiers
CDS 1..1956
tag= a
product= "cp7033"

W0200202606-A2.
PD

10-JAN-2002.
XX PF

03-JUL-2001; 2001WO-IB001445.
XX PF

03-JUL-2000; 2000GB-00016363.
PR PR

11-JUL-2000; 2000GB-00017047.
PR PR

21-JUL-2000; 2000GB-00017983.
PR PR

07-AUG-2000; 2000GB-00019368.
PR PR

18-AUG-2000; 2000GB-00020440.
PR PR

14-SEP-2000; 2000GB-00022583.
PR PR

10-NOV-2000; 2000GB-00027549.
PR PR

22-DEC-2000; 2000GB-00031706.
PR PR

(CHIRP) CHIRON SPA.
PA

Ratti G, Grandi G;
PI PI

WPI; 2002-154726/20.
XX DR

N-PSDB; ABB90532.
XX DR

Novel Chlamydia pneumoniae protein useful in the manufacture of a
medicament for treatment or prevention of infection due to Chlamydia,
preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX FT

Claim 5; Page 47-48; 364pp; English.
XX PS

Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
pneumoniae (strain CML029), and ABL91184-ABL91373 represent DNA encoding
them. The proteins are predicted to be immunogenic and may therefore be
useful in vaccine production and for diagnostic purposes. Chlamydia
pneumoniae is a common cause of respiratory disease in humans, and is
also involved in the development of cardiovascular diseases such as
atherosclerosis, coronary artery disease, carotid artery stenosis,
myocardial infarction, cerebrovascular disease, aortic aneurysm,
claudication and stroke. The proteins and nucleic acids of the invention
may be used in vaccines and pharmaceutical compositions for the
prevention or treatment of chlamydial infections, particularly Chlamydia
pneumoniae infections. The proteins may also be used in the detection of
Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
DNA probe assay or blotting techniques for determining Chlamydia
pneumoniae gene expression. The present sequence represents a
specifically claimed DNA which encodes a Chlamydia pneumoniae protein of
the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX CC

Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;
SQ

Query Match 93.1%; Score 1355; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

101 ATGGTAAATCTATTGTCAGGTCCATATAGACGAACAAGCGCACACTCCGCAGAT 160

xx The present invention describes compounds and methods for diagnosing and
CC treating Chlamydial infection. Chlamydia polynucleotide and protein
CC sequences have antibiotic, anti-inflammatory, antifertility, cardiant,
CC anti-arteriosclerotic and ophthalmological activities, and can be used in
CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
CC compositions or methods from the present invention can be used for the
CC serodiagnosis or methods from the present invention can be used for the
CC serodiagnosis or treatment of Chlamydial infections, particularly in
CC humans. The polynucleotides, proteins or compositions are particularly
CC useful for stimulating an immune response in a patient, or for
CC stimulating and/or expanding T cells specific for a Chlamydia protein.
CC

QY 1085 GCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAGGA 1144
Db 781 GCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAGGA 840
QY 1145 AGTAGATTGGTAGTATCGTGTTTCCATGCTGTTAGATGATCTGAATGAGCGCT 1204
Db 841 AGTAGATTGGTAGTATCGTGTTTCCATGCTGTTAGATGATCTGAATGAGCGCT 900
QY 1205 TCCATTTTGAATGCTGGGTTTCCATGCTGATGATTCACATGTTCAATACGGAAATCCTGAT 1264
Db 901 TCCATTTTGAATGCTGGGTTTCCATGCTGATGATTCACATGTTCAATACGGAAATCCTGAT 960
QY 1265 TCTCAAGTCCCAACAGGAGCTCGGAGCAACAGCTAGAGCAGCAAGCGCTGGAGAT 1324
Db 961 TCTCAAGTCCCAACAGGAGCTCGGAGCAACAGCTAGAGCAGCAAGCGCTGGAGAT 1020
QY 1325 GACAGTCTGCTCAGCGCTGGAGATGCTCAGAGCTTTCAGAGCGCTCTAGGTAAA 1384
Db 1021 GACAGTCTGCTCAGCGCTGGAGATGCTCAGAGCTTTCAGAGCGCTCTAGGTAAA 1080
QY 1385 GCTGGGCAACAAACAGGCACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTG 1444
Db 1081 GCTGGGCAACAAACAGGCACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTG 1140
QY 1445 AGCGCAGGAGT 1455
Db 1141 AGCGCAGGAGT 1151

RESULT 10
AAZ28411
ID AAA28411 standard; DNA; 1550 BP.

AC AAA28411;
DT 15-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)
XX Chlamydia pneumoniae lorf2 coding sequence.
XX Lorf2; vaccine; antibacterial; antigen; ss.
XX Chlamydia pneumoniae.

Key Location/Qualifiers
CDS 101..1369
/*tag= a

W0200024901-A1.

04-MAY-2000.

28-OCT-1999; 99WO-GB003565.

28-OCT-1998; 98US-0106037P.

20-SEP-1999; 99US-0154658P.

26-OCT-1999; 99US-00427501.

(CONN-) CONNAUGHT LAB LTD.

Murkin AD, Oomen RP, Dunn PL;

WPI; 2000-350742/30.

P-PSDB; AAY92716.

Isolated polynucleotide encoding a Chlamydia polypeptide useful to treat,

diagnose and prevent disease caused by Chlamydia infection.

Claim 1; Fig 1A-C; 8pp; English.

This sequence encodes lorf2 protein of a strain of Chlamydia pneumoniae.

Comparison of this sequence as to the recently published genome sequence

of C. pneumoniae reveals that the sequence actually contains at least two
open reading frames, a first one in the 5' portion and a second one in
the 3' portion of the sequence. Despite the presence of the stop codon at
the end of this sequence, C. pneumoniae does make a 76 kDa product. It
appears possible that C. pneumoniae is able to read through this stop
codon and produce a full-length product terminated by the stop codon at
the end of the second open reading frame. There is at least one in-frame
Arg upstream of the start codon. This suggests that the first open
reading frame may form part of one or more larger open reading frames.
The lorf2 protein or DNA can be used as a vaccine for humans to treat or
prevent disease caused by Chlamydia infection. The sequences or an
antibody to lorf2 can be used to diagnose a Chlamydial infection.
(Updated on 15-SEP-2003 to standardise OS field)

Sequence 1550 BP; 472 A; 268 C; 317 G; 493 T; 0 U; 0 Other;

Query Match 11.7%; Score 170; DB 3; Length 1550;

Best Local Similarity 100.0%; Pred. No. 4.8e-32;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTAATAGTCAGAGCTTTTATTTT 60

Db 1381 ATAAATCTTTAAACAGGCTCGCATTAATTAATAGTCAGAGCTTTTATTTT 1440

QY 61 ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120

Db 1441 ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1500

QY 121 AGTCTCTATAGACGAACAGACGACACCTCCGCGAGATCTTCTGCTC 170

Db 1501 AGTCTCTATAGACGAACAGACGACACCTCCGCGAGATCTTCTGCTC 1550

RESULT 11

AAV16207

ID AAV16207 standard; DNA; 150 BP.

AC AAV16207;

DT 17-OCT-2003 (revised)

DT 28-MAY-1998 (first entry)

Part of the gene encoding the 76 kDa cysteine rich OMP.

Cysteine rich outer-membrane protein; OMP; 60 kDa OMP; 76 kDa OMP;

hybridisation; amplification; assay; detection; ds.

Chlamydia pneumoniae.

W09746705-A2.

11-DEC-1997.

04-JUN-1997; 97WO-US009673.

06-JUN-1996; 96US-00659473.

(ABBO) ABBOTT LAB.

Cerney MB;

WPI; 1998-042215/04.

Chlamydia pneumoniae derived oligonucleotides - used as primers or probes

for specific and sensitive detection.

Disclosure; Page 24; 28pp; English.

The present sequence represents part of the gene encoding the 76 kDa

cysteine rich outer-membrane protein (OMP) of Chlamydia pneumoniae. Novel

oligonucleotides AAV16195-206 can function either as PCR primers or

probes. They are used to detect C. pneumoniae in a sample.

Oligonucleotides V161202-04 are specific for the present sequence. The

XX (CORI-) CORIXA CORP.
XX Fling SP, Skeiky YAW, Probst P, Bhatia A;
XX WPI; 2002-179901/23.
XX Novel compositions comprising Chlamydia CapI protein and its use in the
XX treatment of Chlamydia infection.
XX Claim 34; Page 374-375; 537pp; English.
XX The present invention describes compositions comprising a Chlamydia CapI
XX protein and methods for the diagnosis and therapy of Chlamydia infection.
XX Chlamydia DNA and protein sequences from the present invention can have
XX antibacterial and immunostimulant activities, and can be used in
XX vaccines. Compounds from the present invention can be used for eliciting
XX an immune response, specifically stimulating a Chlamydia-specific T-cell
XX response or inhibiting the development of a Chlamydia infection in an
XX animal. Methods from the present invention can be used: for detecting the
XX presence of Chlamydia in a patient; to stimulate and/or expand T cells
XX specific for a Chlamydia protein; and for treatment of a Chlamydia
XX infection. ABL92394 to ABL92709 and ABL94096 to ABL94374 represent
XX sequences used in the exemplification of the present invention
XX
SQ Sequence 1941 BP; 610 A; 391 C; 445 G; 495 T; 0 U; 0 Other;

Query Match 3.7%; Score 54.4; DB 6; Length 1941;
Best Local Similarity 45.9%; Pred. No. 0.0038;
Matches 402; Conservative 0; Mismatches 441; Indels 33; Gaps 5;

QY 606 AAGTTGGCGGCAATACAGATTAGCGAAATATGCTTCGGATACCAAGCGATTCTTG 665
DB 467 AGTTAATGATAGCTCAATCAGTTAGTTAACTTGAATAACAGATTCAGACTTTAAAGG 526
QY 666 ACTCTTAGTAACTGACTTCCTTCCAGCTCTTACAGGCTGCTCTTCTCAATCTGTAG 725
DB 527 AACTTTAAACCAACAGACTCTGCAGATCAGATCCAGCGATTAATAGTCAGTAGAGA 586
QY 726 CAACATAACAAGCAGCTGAGCTCTTAAAGAGATGCAAGATAACCCAGTAGTCCAG 785
DB 587 TCAACAAAATCTGCAGATCAAAATATCAAGATCTGGAAGGACAAAACATAAGTTATG 646
QY 786 GGAACGCTGCAATCTTCTTCTTATGATCAGAGATGCTACAGCAGACAGA 845
DB 647 AAGCTGTTCTCAAGCAGGAGAGGTTATCAAGCTTCTTGAAGCGGAATTAAAT 706
QY 846 TAGAGAAAGATGGAATGCGATTAGGATGCATATTTTGCAGGACAGAAAGCTAGTGGAG 905
DB 707 TAGGACAGCTTTGCACTTATTTGATGCTGGGATCAAGCCAGGCTGCAGTTCTTC 766
QY 906 CTGTAGAAATGCTAAATCTAATAACAGTATAGCAATAGATTCAGCTAAAGCAGCAA 965
DB 767 AAGCACAGCAAAATAATAGCCAGATAATATCGCAGCAGCAAGAAATTAATGATGCTG 826
QY 966 TCGCTACTGCTAAGACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCATTC 1025
DB 827 CTGAACGAAGGTAAACGAGTTAAACAGAGATACAGGCTAACGGATCGCCTTTAG 886
QY 1026 TTCAAGAAGCGGAACAAATGGTAATACAGCTGAGAAAGATCTTAAATAATCAAACTG 1085
DB 887 TGAATAAAGCTGAGGAGCAGATTAGTCAAGCACAATAAGATATTCAGAGATCAACCT- 945
QY 1086 CAGATGTTCTGATGTTCCAAATCCAGAACTA--CAGTTGGAGGCTCCAAAGCAACAGG 1143
DB 946 --AGTGTTCGGATATTCCTATCGTTCGGTCCGAGTGGCTCAGCTGCTCCGAGGAAGTG 1003
QY 1144 AAGTAGTATGGTAGTATTC-----GTGTTTCATGCTCTTAGATCATG 1187
DB 1004 CGGTAGGAGCGTTGAATTCCTTAACAAATTCAGGAAGAAATTCCTGCTTGTATGATG 1063
QY 1188 CTGAAAATGAGACCGCTTCCATTTTGTATGCTGGGTTTCGTCAGATGATTCAATGTTCA 1247

DB 1064 TAGACAATGAATGGCAGCGATTGCAATGCAAGGTTTTCGATCTATGATGAAACAATTYA 1123
QY 1248 ATACGGAAAATCCTGATTCTCAAGCTGCCCAACAGAGAGCTCGCAGCAACAGCTAGAGCAG 1307
DB 1124 ATGTAAACAATCTGCAACAGCTAAAGAGCTACAAGCTATGGAGGCTCAGTCGA---CTG 1180
QY 1308 CGAAACCCGCTGAGATGACAGTGTCTGCTGCGAGCGCTGCAGATGCTCAGAAAGCTTTAG 1367
DB 1181 CGATGTGAGATCAATCACTGGTTGGTGGGATGGCGAGCTCCAGCCGAATACAAGCAATCA 1240
QY 1368 AAGCGCTCTAGTAAAGCTGGGCAACAAC-----AGGCATACTCAATGCTTTAG 1418
DB 1241 AAGATGCTCTTGGCAGAGCTTTGAACAACCATCAACAGATGGTTTAGCTACAGCTATGG 1300
QY 1419 GACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAG 1454
DB 1301 GACAAGTGGCTTTTGCGAGCTGCCAAGGTTGGAGGAG 1336

RESULT 14
ADD43867
ID ADD43867 standard; DNA; 1944 BP.
XX AC ADD43867;
XX AC ADD43867;
DT 15-JAN-2004 (first entry)
XX Chlamydia trachomatis immunogenic DNA sequence, SEQ ID No 162.
XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
XX gene therapy; antibacterial; gene; ds.
XX Chlamydia trachomatis.
XX WO2003049762-A2.
XX 19-JUN-2003.
XX 12-DEC-2002; 2002WO-IB005761.
XX 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX (CHIR-) CHIRON SPA.
XX Grandi G, Ratti G;
XX WPI; 2003-532882/50.
DR P-PSDB; ADD43866.
XX New immunogenic composition having a protein or encoding nucleic acid,
XX useful for diagnosing, preventing and/or treating Chlamydia trachomatis
XX infection.
XX Claim 7; SEQ ID NO 162; 164pp; English.
XX The invention relates to a novel immunogenic composition comprising a
XX protein or nucleic acid, and an adjuvant, where the protein or nucleic
XX acid comprises any of 131 fully defined amino acid or nucleotide
XX sequences given in the specification, or has 50% or greater sequence
XX identity to it, or their fragments. The protein and/or nucleic acid of
XX the immunogenic composition is useful in the manufacture of a medicament
XX for the treatment or prevention of infection due to Chlamydia
XX trachomatis. The infection is treated or prevented by the medicament
XX eliciting an immune response which is specific to a C. trachomatis
XX elementary body, or for neutralising C. trachomatis elementary bodies,
XX hence the immunogenic composition can be used in creating a vaccine. The
XX immunogenic compositions can also be used for the diagnosis of C.
XX trachomatis infection. The nucleic acids of the immunogenic compositions
XX can be used to treat disorders by gene therapy. The immunogenic
XX compositions have antibacterial activity. This polynucleotide sequence
XX represents one of the 131 DNA sequences encoding the C. trachomatis

CC proteins with immunogenic properties of the invention.

XX Sequence 1944 BP; 612 A; 391 C; 445 G; 496 T; 0 U; 0 Other;

SQ Query Match 3.7%; Score 54.4; DB 9; Length 1944;
Best Local Similarity 45.9%; Pred. No. 0.0039;
Matches 402; Conservative 0; Mismatches 441; Indels 33; Gaps 5;

QY 606 AAGTTGGCGCAATACAGATTAGCGAATATGCTTCGGATAACCAACGAGTTCTTG 665
DB |||||
QY 666 ACTCTTTAGTAACTGCTTCCTCGACCTTTACAGGCTGCTTCTTCCAACTGTAG 725
DB |||||
QY 527 AAACCTTTAAACACACAGACTCTGCAGATCAGATTCCAGCGATTAAATAGTCAGTAGAGA 586
DB |||||
QY 726 CAAACAAATACAAACAGCTGAGCTCTTAAAGAGATCAAGATAACCCAGTAGTCCCGAC 785
DB |||||
QY 587 TCAACAAAATTTCTGCAGATCAAAATATCAAGATCTGGAAGGACAAACATTAATTATG 646
DB |||||
QY 786 GGAAACGCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAGA 845
DB |||||
QY 647 AAGCTGTTCTACTAACGACGAGAGGTTATCABAGCTTCTTCTGAAGCGGGAATTAAGT 706
DB |||||
QY 846 TAGAGAAAGATGGAATGCGATTAGGATGCAATTTTTCAGGACAGACGCTAGTGGAG 905
DB |||||
QY 707 TAGGACAAAGCTTTGCACTATTTGATGCTGCGGATCAAGCCAGGCTGCAGTCTTTC 766
DB |||||
QY 906 CTGTAGAAATGCTAAATCTAATACAGTATTAAGCAACATAGATTTCAGCTAAAGCAGGAA 965
DB |||||
QY 767 AAGCAGCAAAATATATAGCCAGATATATCGAGCCACAGAAATTAATTGATGCTG 826
DB |||||
QY 966 TCGCTACTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAAATTC 1025
DB |||||
QY 827 CTGAACGAGGTTAAACGAGTTAAACAAAGAGCATACAGGCTAACCGACTCGCTTTAG 886
DB |||||
QY 1026 TTCAGAACGGGAAACAAATGTTAATACAGGCTGAGAAAGATCTTAAATATCAACCTG 1085
DB |||||
QY 887 TGAATAAAGCTCAGGAGCAGATTAGTCAAGSCACAAAGATATTCAGAGATCAAAACCTT- 945
DB |||||
QY 1086 CAGATGTTCTGATGTTCCAAATCCAGGAACCTA- -CAGTTGGAGGCTCCAAAGCAACAGG 1143
DB |||||
QY 946 --AGTGGTTCGGATATCTTATCGTTGGTCCGAGTGGTCACTGTTCCGAGGAGTG 1003
DB |||||
QY 1144 AAGTAGTATTGGTAGTATTC- - - - -GTGTTCCATGCTGTTAGATGATG 1187
DB |||||
QY 1004 CGGTAGGAGGTTGAAATCTCTACAAATTCAGGAAGAAATTCCTTGTGCTGATGATG 1063
DB |||||
QY 1188 CTGAAATGAGACCGCTTCCATTTTGTCTGCTGGTTTCGTACAGATGATTCACATGTTCA 1247
DB |||||
QY 1064 TAGACAAATGAATGGCAGCGATTGCAATGCAAGGTTTGTGATCTATGATCGAACAAATTA 1123
DB |||||
QY 1248 ATACGGAATATCTGATTTCTCAAGCTGCCCAACAGGAGCTTCGAGCACAAAGCTTAGAGCAG 1307
DB |||||
QY 1124 ATGTAACATCTCTGCAACAGCTTAAGAGCTACAGCTATGGAGGCTCAGCTGA- - -CTG 1180
DB |||||
QY 1308 CGAAGCCGCTGGAGATGACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAG 1367
DB |||||
QY 1181 CGATGTCAGATCACTGTTGTTGGTGGGATGGGAGTCCCGAGGCTCCCGAGGAAATCAAGCAATCA 1240
DB |||||
QY 1368 AAGGGCTCTAGTAAAGCTGGGCAACAC- - - - -AGGCATATCAATCAATGCTTTAG 1418
DB |||||
QY 1241 AAGATGCTCTTGGCAAGCTTTGAAACAAACCATCAACAGATGTTTGTAGCTACAGCTATGG 1300
DB |||||
QY 1419 GACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAG 1454
DB |||||
QY 1301 GACAAAGTGGCTTTGGAGCTGCCAAGGTTGGAGGAG 1336
DB |||||

RESULT 15
AAZ01425_07
Continuation (8 of 11) of AAZ01425 from base 700001 (Complete genome sequence of Chlamydia
WP Sequence split into 11 fragments LOCUS AAZ01425 Accession Aaz01425

WP Fragment Name Begin End
AAZ01425_00 1 110000
WP AAZ01425_01 100001 210000
WP AAZ01425_02 200001 310000
WP AAZ01425_03 300001 410000
WP AAZ01425_04 400001 510000
WP AAZ01425_05 500001 610000
WP AAZ01425_06 600001 710000
WP AAZ01425_07 700001 810000
WP AAZ01425_08 800001 910000
WP AAZ01425_09 900001 1010000
WP AAZ01425_10 1000001 1100000

Query Match 3.6%; Score 51.8; DB 2; Length 110000;
Best Local Similarity 50.8%; Pred. No. 0.093;
Matches 248; Conservative 0; Mismatches 208; Indels 32; Gaps 4;
QY 993 AAGCTCAGAAAAAGTTCCCGACTCTCCAATTTTCAAGAGCGGAACAATGTTATATAC 1052
DB |||||
QY 105190 AAGAACATCAAGAAATAGCGACTCGCTTTAGTGAATAAAGCTGAGGAGCAGATTAAATC 105249
DB |||||
QY 1053 AGGCTGAGAAAGATCTTAAATAATATCAACTGCAGATGTTCTGATGTTCCCAATCCAG 1112
DB |||||
QY 105250 AAGCACAAACAGATATTCAACAGATCACT- - -AGTGGTTTGGATATTCCTATCGTTG 105306
DB |||||
QY 1113 GAACTACAGTTGGAGGCTCCAAAGCAACAAAGGAAGTAGTATTGGTAGTATTC- - - - - 1163
DB |||||
QY 105307 GTCCGATGGGTGAGGTTTCCGAGGAGTGGCGGAGGCGTTGAAATCCTCTAACAA 105366
DB |||||
QY 1164 - - - - -GTGTTTCCATGCTGTTAGATGATGCTGGAATAAGTACGCGCTTCCATTTTAT 1215
DB |||||
QY 105367 TTCAGGAAGAAATTTCTTGTGTTGATGATGATGATGATGATGATGATGATGATGATG 105426
DB |||||
QY 1216 GTCTGGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1275
DB |||||
QY 105427 GCAAGGTTTTCGATCTATGATCGAACAAATTTAATGTAACAAATCTCTGCAACAGCTAAAGA 105486
DB |||||
QY 1276 CCAACAGGAGCTCGCAGCACCAAGCTTAGAGCAGCGAAGCCGCTGGAGATGACAGTGTGC 1335
DB |||||
QY 105487 GCTACAGCTATGGAGGCTCAGCTGA- - -CTGCGATGTCAGATCACTGTTGGTGGCGA 105543
DB |||||
QY 1336 TGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAACGGCTCTAGTAAAGCTGGGCAACA 1395
DB |||||
QY 105544 TGGCAGGCTCCAGCGGAAATACAGCAATCAAGATGCTCTTGGCAGGCTTTGAACA 105603
DB |||||
QY 1396 AC- - - - -AGGCATATCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTGAG 1446
DB |||||
QY 105604 ACCATCAGCAGATGTTTGTAGCTACAGCTATGGCAACAGTGGCTTTTGCAGCTGCCAAGCT 105663
DB |||||
QY 1447 CGCAGGAG 1454
DB |||||
QY 105664 TGGAGGAG 105671
DB |||||

Search completed: March 23, 2004, 20:03:04
Job time : 563.656 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:22:08 ; Search time 100.993 Seconds
(without alignments)
8000.664 Million cell updates/sec

Title: US-10-608-559-5
Perfect score: 1456
Sequence: 1 ataaatctttaaaacagg.....ctgttgagcaggaagta 1456

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1444	99.2	1230025	4	US-09-198-452A-1
2	170	11.7	1550	4	US-09-427-501-1
3	170	11.7	1550	4	US-09-905-119A-1
4	148.4	10.2	150	3	US-08-659-473-8
5	54.4	3.7	7218	1	US-08-232-463-14
6	51.6	3.5	2205	4	US-09-328-352-286
7	51.4	3.5	1537	4	US-09-556-877-57
8	51.4	3.5	1537	4	US-09-620-412C-57
9	51.4	3.5	1537	4	US-09-410-568-57
10	51.4	3.5	1537	4	US-09-598-419-57
11	49	3.4	1990	4	US-08-714-741-35
12	49	3.4	2085	1	US-08-072-070-1
13	49	3.4	2085	1	US-08-465-746-1
14	49	3.4	2085	1	US-08-214-164-1
15	49	3.4	2085	1	US-08-469-434-1
16	49	3.4	2085	1	US-08-214-222-1
17	49	3.4	2085	2	US-08-467-852A-1
18	49	3.4	2085	2	US-08-468-719-1
19	49	3.4	2085	2	US-08-246-636-1
20	49	3.4	2085	2	US-08-247-491A-1
21	49	3.4	2085	2	US-08-319-795-1
22	49	3.4	2085	2	US-08-468-985-1
23	49	3.4	2086	3	US-08-312-949-1
24	49	3.4	2086	3	US-08-446-201-2
25	47.4	3.3	956	4	US-08-714-741-36
26	43.6	3.0	6744	1	US-08-119-125A-2
27	39.8	2.7	1295	4	US-09-286-981B-20

28	39.8	2.7	7953	4	US-08-956-171E-152
29	39.6	2.7	2208	4	US-09-134-001C-603
30	39.2	2.7	1230025	4	US-09-198-452A-1
C 31	39	2.7	2550	6	5258287-23
C 32	38.8	2.7	1866	4	US-09-601-198-153
33	38.6	2.7	8961	4	US-10-204-708-80
34	38.4	2.6	1196	4	US-08-956-171E-71
35	38.4	2.6	2598	4	US-09-612-204B-23
C 36	38.4	2.6	1664976	4	US-08-916-421B-1
37	38.2	2.6	1254	4	US-09-540-236-547
38	38	2.6	471	4	US-09-134-001C-2415
39	38	2.6	640	3	US-08-961-083-33
40	38	2.6	640	4	US-09-536-784-33
41	38	2.6	6314	4	US-09-620-312D-98
42	38	2.6	6583	4	US-10-204-708-26
43	38	2.6	30549	4	US-09-134-001C-322
44	37.8	2.6	1470	4	US-09-134-001C-2404
C 45	37.8	2.6	3337	1	US-08-072-610-1

ALIGNMENTS

RESULT 1

US-09-198-452A-1/c
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1
; LENGTH: 1230025
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(15000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (15001)..(30000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (30001)..(45000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (45001)..(60000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (60001)..(75000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (75001)..(90000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (90001)..(105000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (105001)..(120000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (120001)..(135000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (135001)..(150000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (150001)..(165000)


```
; LOCATION: (895001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature

Query Match      99.2%; Score 1444; DB 4; Length 1230025;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATAAATCTTTAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 60
Db 828698 ATAAATCTTTAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 828639

QY 61 ATAAATCTTTAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 120
Db 828638 ATAAATCTTTAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 828579

QY 121 AGTCTTATAGCAAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 180
Db 828578 AGTCTTATAGCAAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 828519

QY 181 GCGAGTGCAGCAAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 240
Db 828518 GCGAGTGCAGCAAAACAGAGTCCATTAATATTAGTGCAGAGCTTTTATATTTT 828459

QY 241 TAAAGATCTTAAGACCGATCTCTAGAGCGATGAGAGTCTTGGTCTTCTCAAGTATGC 300
Db 828458 TAAAGATCTTAAGACCGATCTCTAGAGCGATGAGAGTCTTGGTCTTCTCAAGTATGC 828399

QY 301 TCTCATGAGTCTGCAGATAGCTGGGTATGCTTCTAGTAAACAGCTCTTCTACTAG 360
Db 828398 TCTCATGAGTCTGCAGATAGCTGGGTATGCTTCTAGTAAACAGCTCTTCTACTAG 828339

QY 361 CAGATCTGCAGAGTGCAGTCAACAGCAGCGACCGACCTAGGCTCTCTCCACCGATT 420
Db 828338 CAGATCTGCAGAGTGCAGTCAACAGCAGCGACCGACCTAGGCTCTCTCCACCGATT 828279

QY 421 TGATGATTAAGAATCAAGCGCAAAACAGCTTACGATPACTATCTTTACCTCAACATCACT 480
Db 828278 TGATGATTAAGAATCAAGCGCAAAACAGCTTACGATPACTATCTTTACCTCAACATCACT 828219

QY 481 AGTGCATACAGGCTCTTGTGAGCGCTCCAGGATGCTGTCACTAATAAAGATAC 540
Db 828218 AGTGCATACAGGCTCTTGTGAGCGCTCCAGGATGCTGTCACTAATAAAGATAC 828159

QY 541 AGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAACTAAGAAATGCCGATGC 600
Db 828158 AGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAACTAAGAAATGCCGATGC 828099

QY 601 AGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 660
Db 828098 AGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 828039

QY 661 TCTTGATCTTTAGTAAATCTGACTTCTCTCGACCTCTTACAGCTCTCTTCTCAATC 720
Db 828038 TCTTGATCTTTAGTAAATCTGACTTCTCTCGACCTCTTACAGCTCTCTTCTCAATC 827979

QY 721 TGATGCAAAACAAATACAGAGTGCAGTCTTCTTAAAGAGATGCAAGATTAACCCAGTAGT 780
Db 827978 TGATGCAAAACAAATACAGAGTGCAGTCTTCTTAAAGAGATGCAAGATTAACCCAGTAGT 827919

QY 781 CCCAGGAAACCGCTCGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 827918 CCCAGGAAACCGCTCGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 827859

QY 841 ACAGATAGAGAAAGATGGAATGCGATAGGGATGCAATTTTCCAGGACAGAACCGTAG 900
Db 827858 ACAGATAGAGAAAGATGGAATGCGATAGGGATGCAATTTTCCAGGACAGAACCGTAG 827799

QY 901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTTCAGCTAAAGC 960
Db 827798 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTTCAGCTAAAGC 827739

QY 961 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
Db 827738 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 827679

QY 1021 AATCTTCAAGAACCGGAACAATGTAATACAGGCTGAGAAAGATCTTAAATATATCAA 1080
Db 827678 AATCTTCAAGAACCGGAACAATGTAATACAGGCTGAGAAAGATCTTAAATATATCAA 827619

QY 1081 ACCTGCAGATGGTCTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACA 1140
Db 827618 ACCTGCAGATGGTCTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACA 827559

QY 1141 AGGAATAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAATGAGAC 1200
Db 827558 AGGAATAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAATGAGAC 827499

QY 1201 CGCTTCCATTTGATGCTGGGTTTCGTAGATGATTCATGTTCAATACGAAAATCC 1260
Db 827498 CGCTTCCATTTGATGCTGGGTTTCGTAGATGATTCATGTTCAATACGAAAATCC 827439

QY 1261 TGATTTCAAGCTGCCCAACAGAGAGCTCCGAGCACCAAGCTAGAGCAGCGAAAGCGCTGG 1320
Db 827438 TGATTTCAAGCTGCCCAACAGAGAGCTCCGAGCACCAAGCTAGAGCAGCGAAAGCGCTGG 827379

QY 1321 AGATGACGTGCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTTAGG 1380
Db 827378 AGATGACGTGCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTTAGG 827319

QY 1381 TAAAGCTGGCAACAAACAGGAGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTG- 1439
Db 827318 TAAAGCTGGCAACAAACAGGAGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTG- 827259

QY 1440 TTCTGAGCGCAGGAGT 1455
Db 827258 TTCTGAGCGCAGGAGT 827243

RESULT 2
US-09-427-501-1
; Sequence 1, Application US/09427501
; Patent No. 6403101
; GENERAL INFORMATION:
; APPLICANT: Connaught Laboratories Limited, Andrew
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 19721-010
; CURRENT APPLICATION NUMBER: US/09/427,501
; EARLIER FILING DATE: 1999-10-26
; EARLIER APPLICATION NUMBER: 60/106,037
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/154,658
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1366)
US-09-427-501-1

Query Match      11.7%; Score 170; DB 4; Length 1550;
Best Local Similarity 100.0%; Pred. No. 3.7e-36;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCAATTAATTATTAGTGCAGAGCTTTTATATTTT 60
Db 1381 ATAAATCTTTAAACAGGCTCGCAATTAATTATTAGTGCAGAGCTTTTATATTTT 1440
```



```
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match          3.7%; Score 54.4; DB 1; Length 7218;
Best Local Similarity 9.3%; Pred. No. 0.00017;
Matches 43; Conservative 218; Mismatches 199; Indels 0; Gaps 0;

Qy 731 AATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCGAGGAAA 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1481 AATTACCTATCTATGCAAGTAGTTAAAGAGATAGAAAGATTGCTACRRRRRRRRRRR 1422

Qy 791 ACGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAGATAGAG 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1421 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1362

Qy 851 AAAGATGAATGCGATTAGGATGTCATATTTCGAGACAGAAAGCTAGTGGAGCTGTA 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1361 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1302

Qy 911 GAAATGCTAAATCTAATAACATAGTAACCAATAGATTGCTGTAAGACGCAATCGCT 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1301 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1242

Qy 971 ACTGCTAAGACAAATAGCTGAGCTCAGAAAGAGTCCCGACTCTCCAATCTTCAA 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1241 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1182

Qy 1031 GAAGCGCAACAAATGGTAAATACAGCTGAGAAAGATCTTAAATAATCAAACTCGAGAT 1090
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1181 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1122

Qy 1091 GGTCTGTGATTTCCAAATCCAGGAACACTAGCTGGAGGCTCCAGCAACAAAGGAAGTAGT 1150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1121 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1062

Qy 1151 ATTGTAGTATTTCGGTTTCATGCTGCTGTAGATGATGCTG 1190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1061 AAGCTCCCTCGACCTGAGCCAAAGCTCGGAATTAATCTG 1022

RESULT 6
US-09-328-352-286
; Sequence 286, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 286
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-286

Query Match          3.5%; Score 51.6; DB 4; Length 2205;
```

```
Best Local Similarity 54.8%; Pred. No. 0.00052;
Matches 102; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 791 ACGCCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAGATAGAG 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1426 ACAATTCTATTTCACCCAGCTTTAAACGGAATAAATAACATGCTTCGGTATCAGCTATAGAT 1485

Qy 851 AAAGATGGAATGCGATTAGGATGTCATATTTTCAGGACAGAAACGCTAGTGGAGCTGTA 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1486 AATGCTGGAATAAGTCTGAAGTTGTTGATTTAGGTACAAAGAGATACAACACCAACCA 1545

Qy 911 GAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGCTAAAGCAGCAATCGCT 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1546 GCAAAACCTATATTAATAGCTAGATGATGATGCTGTTAAAGAGGCTATAACG 1605

Qy 971 ACTGCT 976
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1606 GCTGGT 1611

RESULT 7
US-09-556-877-57
; Sequence 57, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 57
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
; US-09-556-877-57

Query Match          3.5%; Score 51.4; DB 4; Length 1537;
Best Local Similarity 53.5%; Pred. No. 0.00049;
Matches 159; Conservative 0; Mismatches 126; Indels 12; Gaps 2;

Qy 1167 TTTCATGCTGTTAGATGATGCTGAAATGAGACCGCTTCCATTTTGTCTGGGTTTC 1226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 TTTCCTTTGCTTGTGATGATGATGACAAATGAAATGGCAGGATGCAATGCAAGGTTTC 245

Qy 1227 GTCAGATGATTCATGCTTCAATACGGAAATCTGATTTCAAGCTGCCCAACAGGAGC 1286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GATCTATGATGAAACAATTTAATGTAACAATCTCGAACAGCTTAAAGAGAGCTACAAGCTA 305

Qy 1287 TCGCAGACAGCTAGAGCAGCAAGCGCTGGAGATGACAGTGTCTCAGCGCTGG 1346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 TGGAGGCTCAGCTGA---CTGCGATGTCAGATCAACTGTTGTCGGATGGCGAGCTCC 362

Qy 1347 CAGATGCTCAGAAAGCTTTAGAACGGCTCTAGTAAAGCTGGGCAACAAC- 1397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CAGCGAAATACAGCAATCAAGATGCTCTTGGCAAGCTTTTGAACAACCAATCAGCAG 422

Qy 1398 AGGCATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGTGTGTGAGCGCAGGAG 1454
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 ATGTTTATGCTACAGCTATGCGACAGTGGCTTTTGCAGCTGCCAAGGTTGGAGGAG 479

RESULT 8
US-09-620-412C-57
; Sequence 57, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
```

Db	246	GATCTATGATCGACAAATTATGTAAACAACTCTCAACAGCTAAAGAGCTACAAGCTA	305
Qy	1287	TGCAGCACAAGCTTAGAGCAGCGAAGCCGCTGGAGATGACAGTCTGCTGTGACGCGTGG	1346
Db	306	TGAGGAGCTCAGCTGA---CTGGCATCTCAGATCAACTGTTGTTGGGATGGCGAGCTCC	362
Qy	1347	CAGATCCTCAGAAAGCTTTAGAAAGCGGCTCTAGTAAAGCTCGGGCAACAC	1397
Db	363	CAGCCGAATACAAGCAATCAAGATGCTCTTGGCAAGCTTTGAACAACCATCAGCAG	422
Qy	1398	AGGGCATCTCAATGTTTAGGACAGATCGTCTTCTGCTGTTGTGAGCGCAGGAG	1454
Db	423	ATGGTTTAGCTACAGCTATGGACAAGTGGCTTTTGAGCTGCGCAAGTTTGGAGGAG	479

```

RESULT 10
US-09-598-419-57
; Sequence 57, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 57
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-598-419-57

```

Qy	1167	TTTCCATGCTTTAGATGATGCTGAAATCAGACCGCTTTCATTTTGATGTCCTGGGTTTC	12286
Db	186	TTTCCCTGTTGCTTGATGATGTAGCAATGAAATGGCAGGATTTGCATTCGAGTTTTC	245
Qy	1227	GTCAAGATGATTCATGTTTCAATACGGAAATCCTGATTTCTCAAGTCGCCCAACAGGAGC	12866
Db	246	GATCTATGTCGAACAATTTAATGTAACAATCTGCAACAGCTTAAAGAGGCTACAAGCTA	305
Qy	1287	TGCAGCACACAGCTTAGAGCAGCGAAGCCGCTGGAGATGACAGTGCTGCTGCAGCGCTGG	13486
Db	306	TGGAGGCTCAGCTCA---CTTGGCATGTCAATCAACTGTTTGGTTCGGATGGCGAGCTCC	362
Qy	1347	CAGATGCTCAGAAGCTTTAGACGGGCTCTAGTAAAGCTGGCGCAACAC	13976
Db	363	CAGCCGAAATACAGCAATCAAGAATCTCTTGGCAAGCTTTGAAACAACCATCAGCAG	422
Qy	1398	AGGCATACTCAATGCTTTAGGCAGATCGCTTCGTGCTGTTGTGTGAGCGCAGGAG	1454
Db	423	ATGCTTTAGCTACAGCTATCGGACAAGTGGCTTTTTCAGCTGCGCAAGTTGAGGAG	479

Patent No. 6500513
GENERAL INFORMATION:
APPLICANT: Briles, David E.
APPLICANT: McDaniell, Larry S.
APPLICANT: Swiatlo, Edwin
APPLICANT: Yother, Janet
APPLICANT: Crain, Marilyn J.
APPLICANT: Hollingshead, Susan
APPLICANT: Tart, Rebecca

APPLICANT: Brooks-Walter, Alexis
TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
TITLE OF INVENTION: PORTIONS AND PRODUCTS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,741
FILING DATE: 16-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer Esq., William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454312-2460
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1990 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-714-741-35

Query Match 3.4%; Score 49; DB 4; Length 1990;
Best Local Similarity 49.5%; Pred. No. 0.0025;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy 845 ATAGAGAAAGATGGAATGCGATAGGATGCATATTTTCAGGACAGACGCTAGTGGA 904
Db |||||
472 ACAGACAAAGCGCAAGAGCCAGCAGATAGATGATGATGAGCTAAGAAACGGAA 531
Qy 905 GCTGTAGAAATGCTTAATCTAATTAACAGTATAGCAACATAGATTCTAGCTAAGACGCA 964
Db |||||
532 GAAGAGGCAAAACCTAAATTTAATCTGTCGACCAATGGTAGTCTCTGAGCCAGACG 591
Qy 965 ATCGCTACTGTAAAGACAAATAGCTGAAGCTCAGAAAAAGTTCCCGA--CTCTCCA 1021
Db |||||
592 TTGCTGAGCTAGAGAAATACAGAGAGCTTAACAAAAGACACCACTTACTTAA 651
Qy 1022 ATCTTCAAGAGCGGACAAATAGTATACAGCTGAGAAGATCTTTAAAAATATCAAA 1081
Db 652 AAACCTAGAAGAGCTAAAGCAAAATTAAGAGGCTGAGAAAAAGCTACTGAAGCCAAA 711
Qy 1082 CTGCAAGTGTCTGATGTTCCAAATCCAGAACTACAGTTGAGGGTCCCAAGCAAA 1141
Db 712 CAAAAAGTGGATGCTGAGAGAGTCGCTCTCAAGCTAAATCGTGAATTCGGAATCA 771
Qy 1142 GGAAGTAGTATTG 1154
Db 772 GTTCATAGACTAG 784

RESULT 12

US-08-072-070-1
Sequence 1, Application US/08072070
Patent No. 5476929
GENERAL INFORMATION:
APPLICANT: Briles, David E
APPLICANT: Yother, Janet L

APPLICANT: McDaniel, Larry S
TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd
STREET: Suite 1203, 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,070
FILING DATE: 19930603
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/835,698
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/656,773
FILING DATE: 15-FEB-1991
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 521-0378
TELEX: LUKPAT WASHINGTON
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2085 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: Double
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: Rx1
IMMEDIATE SOURCE:
CLONE: JY2008
FEATURE:
NAME/KEY: intron
LOCATION: 1..2085
FEATURE:
NAME/KEY: CDS
LOCATION: join(127..1983, 1987..1992, 1996..2007, 2011
LOCATION: ..2025, 2029..2031, 2035..2085)
US-08-072-070-1

Query Match 3.4%; Score 49; DB 1; Length 2085;
Best Local Similarity 49.5%; Pred. No. 0.0025;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy 845 ATAGAGAAAGATGGAATGCGATAGGATGCATATTTTCAGGACAGACGCTAGTGGA 904
Db |||||
472 ACAGACAAAGCGCAAGAGCCAGCAGATAGATGATGATGAGCTAAGAAACGGCA 531
Qy 905 GCTGTAGAAATGCTTAATCTAATTAACAGTATAGCAACATAGATTCTAGCTAAGACGCA 964
Db |||||
532 GAAGAGGCAAAACCTAAATTTAATCTGTTTCGAGCAATGGTAGTCTCTGAGCCAGACG 591
Qy 965 ATCGCTACTGTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGA--CTCTCCA 1021
Db 592 TTGCTGAGACTAAGAAAAAATCAGAGAGCTAAACAAAAGACCACTTACTTAA 651
Qy 1022 ATCTTCAAGAGCGGACAAATGGTATACAGCTGAGAAGATCTTTAAAAATATCAAA 1081
Db 652 AAACCTAGAAGAGCTAAGCAAAATTAAGAGAGGCTGAGAAAAAGCTACTGAGCCAAA 711

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:23:54 ; Search time 502.565 Seconds
(without alignments)
10716.954 Million cell updates/sec

Title: US-10-608-559-5

Perfect score: 1456

Sequence: 1 ataaatctttaaaacagg.....ctgttgagcaggagta 1456

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubna/PCNT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubna/PCNT_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1456	100.0	1456	10	US-09-564-479-5	Sequence 5, Appli
2	1456	100.0	2238	10	US-09-564-479-7	Sequence 7, Appli
3	1455	99.9	2156	10	US-09-564-479-4	Sequence 1, Appli
4	1444	99.2	1230025	15	US-10-289-762-1	Sequence 1, Appli
5	1355	93.1	1956	9	US-09-841-132-385	Sequence 385, App
6	1355	93.1	1956	10	US-09-841-260-63	Sequence 63, Appl
7	1355	93.1	1956	10	US-10-007-693-63	Sequence 63, Appl
8	1355	93.1	1956	15	US-10-312-273-14	Sequence 14, Appl
9	1151	79.1	1852	10	US-09-564-479-3	Sequence 3, Appli
10	1170	11.7	1550	9	US-09-905-119-1	Sequence 1, Appli
11	54.4	3.7	1941	9	US-09-841-132-412	Sequence 412, App
12	51.4	3.5	1537	9	US-09-841-132-57	Sequence 57, Appl
13	49	3.4	1860	12	US-10-282-122A-37518	Sequence 37518, A
14	48.6	3.3	1171	10	US-09-841-260-13	Sequence 13, Appl
15	48.6	3.3	1171	13	US-10-007-693-13	Sequence 13, Appl

16	48.6	3.3	1834	10	US-09-841-260-38	Sequence 38, Appl
17	48.6	3.3	1834	13	US-10-007-693-38	Sequence 38, Appl
18	48.6	3.3	1983	10	US-09-841-260-28	Sequence 28, Appl
19	48.6	3.3	1983	13	US-10-007-693-28	Sequence 28, Appl
20	47	3.2	9373	12	US-10-221-613-178	Sequence 178, App
21	47	3.2	9373	14	US-10-311-455-1264	Sequence 1264, Ap
22	44.4	3.0	2499	9	US-09-842-552-92	Sequence 92, Appl
23	44.2	3.0	4997	12	US-10-282-122A-35506	Sequence 35506, A
24	44	3.0	3673778	14	US-10-312-841-2	Sequence 2, Appli
25	43.8	3.0	702	15	US-10-027-632-45880	Sequence 45880, A
26	43.8	3.0	702	15	US-10-027-632-45881	Sequence 45881, A
27	43.5	3.0	6341	14	US-10-311-455-1617	Sequence 1617, Ap
28	43	3.0	18218	14	US-10-311-455-1922	Sequence 1922, Ap
29	43	3.0	18585	14	US-10-240-485-161	Sequence 161, App
30	43	3.0	3673778	14	US-10-312-841-1	Sequence 1, Appli
31	42.8	2.9	302	12	US-10-424-599-63444	Sequence 63444, A
32	42.8	2.9	2404	15	US-10-104-047-1740	Sequence 1740, Ap
33	42.8	2.9	3439	15	US-10-161-927-53	Sequence 53, Appl
34	42.8	2.9	16688	14	US-10-311-455-293	Sequence 293, App
35	42.6	2.9	6126	14	US-10-311-455-1418	Sequence 1418, Ap
36	42.4	2.9	330	10	US-09-814-353-5359	Sequence 5359, Ap
37	42.4	2.9	330	10	US-09-814-353-11646	Sequence 11646, A
38	42.4	2.9	5503	16	US-10-257-166-104	Sequence 104, App
39	42.2	2.9	6692	12	US-10-221-714A-131	Sequence 131, App
40	42	2.9	6535	14	US-10-311-455-909	Sequence 909, App
41	41.8	2.9	5278	12	US-10-221-714A-97	Sequence 97, Appl
42	41.8	2.9	5278	14	US-10-311-455-795	Sequence 795, App
43	41.8	2.9	6338	14	US-10-311-455-1099	Sequence 1099, Ap
44	41.6	2.9	488	10	US-09-918-995-31590	Sequence 31590, A
45	41.6	2.9	576	10	US-09-814-353-5208	Sequence 5208, Ap

ALIGNMENTS

RESULT 1
US-09-564-479-5
; Sequence 5, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: US\$ THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 1993-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1456
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1456)
US-09-564-479-5

Query Match	100.0%;	Score 1456;	DB 10;	Length 1456;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1456;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATAAAATCTTTAAACAGCGCTGCATTATTTATTAGTAGAGAGCTTTTATTTT	60	
Db	1	ATAAAATCTTTAAACAGCGCTGCATTATTTATTAGTAGAGAGCTTTTATTTT	60	
Qy	61	ATAATAAACTAAAGATTTTATTTTATTTTATGTTTATGTTTATCTTATGTTCC	120	

Db 906 TAAAGAACTTAAGACCGATTCGTAGAGCGATGGAGCATCTTGGCTTCGAGTGAATGC 965
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTACAGCTCGTCTTCTACTAG 360
Db 966 TCTCATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTACAGCTCGTCTTCTACTAG 1025
QY 361 CAGATCTGAGACGTGGACTCAACGACGCGACCGCTTCCGCTCTCCACCCACGTT 420
Db 1026 CAGATCTGAGACGTGGACTCAACGACGCGACCGCTTCCGCTCTCCACCCACGTT 1085
QY 421 TCATGATTTATAGACTCAAGCGCAACACGCTTACGATCTATCTTTACCTCAACATCACT 480
Db 1086 TCATGATTTATAGACTCAAGCGCAACACGCTTACGATCTATCTTTACCTCAACATCACT 1145
QY 481 AGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCATCAATATAAGGATAC 540
Db 1146 AGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGCATCAATATAAGGATAC 1205
QY 541 AGCGGCTACTGATGAGGAACCGGAATCGCTGGGAGTGGGAAACTAAGAAATGCGGATGC 600
Db 1206 AGCGGCTACTGATGAGGAACCGGAATCGCTGGGAGTGGGAAACTAAGAAATGCGGATGC 1265
QY 601 AGTTAAAGTTGGCGCAATACAGAAATAGCGAATATGCTTCGGATAACCAAGCGAT 660
Db 1266 AGTTAAAGTTGGCGCAATACAGAAATAGCGAATATGCTTCGGATAACCAAGCGAT 1325
QY 661 TCTTGACTCTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 720
Db 1326 TCTTGACTCTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 1385
QY 721 TGTAGCAACATATAAGAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCGAGTAGT 780
Db 1386 TGTAGCAACATATAAGAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCGAGTAGT 1445
QY 781 CCAGGGAACCGCTGCAATTCCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGAC 840
Db 1446 CCAGGGAACCGCTGCAATTCCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGAC 1505
QY 841 ACAGATAGAGAAAGATGGAAATCGATTAGGATGCATATTTTGCAGGACAGAACGCTAG 900
Db 1506 ACAGATAGAGAAAGATGGAAATCGATTAGGATGCATATTTTGCAGGACAGAACGCTAG 1565
QY 901 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC 960
Db 1566 TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGC 1625
QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCC 1020
Db 1626 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCC 1685
QY 1021 AATTTCTCAAGAGCGGAACAAATGTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 1686 AATTTCTCAAGAGCGGAACAAATGTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1745
QY 1081 ACTGAGATGGTCTGATGTTCCCAATCCAGAACTACAGTTGGAGGCTCCAGGCAAC 1140
Db 1746 ACCTGAGATGGTCTGATGTTCCCAATCCAGAACTACAGTTGGAGGCTCCAGGCAAC 1805
QY 1141 AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAATGAGAC 1200
Db 1806 AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAATGAGAC 1865
QY 1201 CGCTTCATTTTGAATGCTGGGTTTCGTAGATGATTCACATGTTCAATACGGAATATCC 1260
Db 1866 CGCTTCATTTTGAATGCTGGGTTTCGTAGATGATTCACATGTTCAATACGGAATATCC 1925
QY 1261 TGAATCTCAGCTGCCCAACAGAGAGCTCGACACAGCTAGACGCGGAACCCGCTGG 1320
Db 1926 TGAATCTCAGCTGCCCAACAGAGAGCTCGACACAGCTAGACGCGGAACCCGCTGG 1985
QY 1321 AGATGACAGTGTGCTGACGCGTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTTAG 1380

Db 1986 AGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 2045
QY 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db 2046 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT 2105
QY 1441 TGTGAGCGCAGAGATA 1456
Db 2106 TGTGAGCGCAGAGATA 2121
RESULT 3
US-09-564-479-1
; Sequence 1, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OOMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(2053)
US-09-564-479-1
Query Match 99.9%; Score 1455; DB 10; Length 2156;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAAAACTTTTAAAAACAGGCTCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 60
Db 1 ATAAAACTTTTAAAAACAGGCTCGCAATTAATTAATTAATTAATTAATTAATTAATTTT 60
QY 61 ATAAATAAATAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 61 ATAAATAAATAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
QY 121 AGGTCTCTATAGAGAAACAGAACAGCACACCTCCCGCAGATCTTTCTGCTCAAGGATTGA 180
Db 121 AGGTCTCTATAGAGAAACAGAACAGCACACCTCCCGCAGATCTTTCTGCTCAAGGATTGA 180
QY 181 GCGAGTGCAGAAATAAGATGCGGAGCTCAAGAGATAGAGATGAGGAGGAGGAGCTAAGCC 240
Db 181 GCGAGTGCAGAAATAAGATGCGGAGCTCAAGAGATAGAGATGAGGAGGAGGAGCTAAGCC 240
QY 241 TAAAGATCTAAGACCGATTCGTAGAGCGATGAGGAGCTTTGCGTCTGCAAGTGAATGC 300
Db 241 TAAAGATCTAAGACCGATTCGTAGAGCGATGAGGAGCTTTGCGTCTGCAAGTGAATGC 300
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
Db 301 TCTCATGAGTCTGGCAGATAAGCTGGTATTCCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
QY 361 CAGATCTGAGACGTGGACTCAACGACGCGACCGGACCTAGCCCTCTCCACCCACGTT 420
Db 361 CAGATCTGAGACGTGGACTCAACGACGCGACCGGACCTAGCCCTCTCCACCCACGTT 420
QY 421 TGATGATTTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT 480

Db 421 TGAATGATTAAGACTCAGCGCAACAGCTTACGATACATCTTTACCTCAACATCACT 480
Qy 481 AGCTGACATACAGGCTGTTTGGTGAGGCTCAGAGTCTCTACTAATATAAAGGATAC 540
Db 481 AGCTGACATACAGGCTGTTTGGTGAGGCTCAGAGTCTCTACTAATATAAAGGATAC 540
Qy 541 AGCGGTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAATTAAGATGCCGATGC 600
Db 541 AGCGGTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAATTAAGATGCCGATGC 600
Qy 601 AGTTAAAGTTCGCGCGCAAAATTAACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Db 601 AGTTAAAGTTCGCGCGCAAAATTAACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Qy 661 TCTTGACTCTTAGTAACTGACTTCTCTGACCTCTTACAGCTCTTACAGGCTGCTCTCCCAATC 720
Db 661 TCTTGACTCTTAGTAACTGACTTCTCTGACCTCTTACAGGCTGCTCTCTCCCAATC 720
Qy 721 TGTAGCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db 721 TGTAGCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
Qy 781 CCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 781 CCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Qy 841 ACAGATAGAGAAAGATGAAATGCAATAGGATGATATTTGAGGACAGACAGCGCTAG 900
Db 841 ACAGATAGAGAAAGATGAAATGCAATAGGATGATATTTGAGGACAGACAGCGCTAG 900
Qy 901 TGGAGCTTAGAGAAATGCTAAATCTTAATACAGTATAAGCAACAPAGATTGAGTAAAGC 960
Db 901 TGGAGCTTAGAGAAATGCTAAATCTTAATACAGTATAAGCAACAPAGATTGAGTAAAGC 960
Qy 961 AGCAATCGTCTGCTAGACACAAATAGCTGAACTCAGAAAGATTTCCCGACTCTCC 1020
Db 961 AGCAATCGTCTGCTAGACACAAATAGCTGAACTCAGAAAGATTTCCCGACTCTCC 1020
Qy 1021 AATTTCTCAAGAACGGCAACAAATGGTAAATCAGGCTGAGAAAGATCTTAAATAATCA 1080
Db 1021 AATTTCTCAAGAACGGCAACAAATGGTAAATCAGGCTGAGAAAGATCTTAAATAATCA 1080
Qy 1081 ACCTCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAAAGCAACA 1140
Db 1081 ACCTCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAAAGCAACA 1140
Qy 1141 AGGAAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTAGATGATGCTGAAATGAGAC 1200
Db 1141 AGGAAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTAGATGATGCTGAAATGAGAC 1200
Qy 1201 CGCTTCCATTTGATGCTGCTGGGTTTCGTGAGATGATTCACATGTTCAATACGGAATATCC 1260
Db 1201 CGCTTCCATTTGATGCTGCTGGGTTTCGTGAGATGATTCACATGTTCAATACGGAATATCC 1260
Qy 1261 TGATTTCAAGCTGCCCAACAGAGCTCCGACCAACAGCTAGACAGCAAGAGCCGCTGG 1320
Db 1261 TGATTTCAAGCTGCCCAACAGAGCTCCGACCAACAGCTAGACAGCAAGAGCCGCTGG 1320
Qy 1321 AGATGACAGTGTGCTGACGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
Db 1321 AGATGACAGTGTGCTGACGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
Qy 1381 TAAAGCTGGGCAACACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 1440
Db 1381 TAAAGCTGGGCAACACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 1440
Qy 1441 TGTGACGGCAGGAGT 1455
Db 1441 TGTGACGGCAGGAGT 1455

US-10-289-762-1/c
Sequence 1, Application US/10289762
Publication No. US20040006218A1
GENERAL INFORMATION:
APPLICANT: Griffais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/10/289,762
CURRENT FILING DATE: 2003-03-27
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 1
LENGTH: 1230025
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(15440)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (15001)..(30000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (30001)..(45000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (45001)..(60000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (60001)..(75000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (75001)..(90000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (90001)..(105000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (105001)..(120000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (120001)..(135000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (135001)..(150000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (150001)..(165000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (165001)..(180000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (180001)..(195000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (195001)..(210000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:

```
NAME/KEY: misc_feature
LOCATION: (210001)..(225000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (225001)..(240000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (240001)..(255000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (255001)..(270000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (270001)..(285000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (285001)..(300000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (300001)..(315000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (315001)..(330000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (330001)..(345000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (345001)..(360000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (360001)..(375000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (375001)..(390000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (390001)..(405000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (405001)..(420000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (420001)..(435000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (435001)..(450000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (450001)..(465000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (465001)..(480000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
```

```
LOCATION: (480001)..(495000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (495001)..(510000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (510001)..(525000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (525001)..(540000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (540001)..(555000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (555001)..(570000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (570001)..(585000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (585001)..(600000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (600001)..(615000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (615001)..(630000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (630001)..(645000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (645001)..(660000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (660001)..(675000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (675001)..(690000)
OTHER INFORMATION: n=a or c or g or t
FEATURE:
```


Query Match 99.2%; Score 1444; DB 15; Length 1230025;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 1 ATAAATCTTTAAAAACAGGCTCGCATTAATTATTAGTAGAGAGCTTTTTTTTTTTTTTTT 60
Db 828698 ATAAATCTTTAAAAACAGGCTCGCATTAATTATTAGTAGAGAGCTTTTTTTTTTTTTTTT 828699
QY 61 ATAAATAAATAAAGATTTTTTATTATTTTTCAGTTTATGTTAACTCCATTTCGTCC 120
Db 828638 ATAAATAAATAAAGATTTTTTATTATTTTTCAGTTTATGTTAACTCCATTTCGTCC 828579
QY 121 AGGTCTATAGAGAAACAGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 180
Db 828578 AGGTCTATAGAGAAACAGACACCTCCCGCAGATCTTTCTGCTCAAGGATTGGA 828519
QY 181 GCGAGTGCAGCAATAAGAGTGGGGAAGCTCAAGATAGCAGGTGCGGAGCTAAGCC 240
```

Db 828518 GCGAGTGCAGCAATAAGAGTGGCGAAGCTCAAGAATAGCAGGTGCGGAAGCTTAAGCC 828459
Qy 241 TAAAGAACTTAAGACCGAATCTGTAGAGCGATGGAGCAATCTTGGCTTCTGAGTGAATGC 300
Db 828458 TAAAGAACTTAAGACCGAATCTGTAGAGCGATGGAGCAATCTTGGCTTCTGAGTGAATGC 828399
Qy 301 TCTCATGAGTCTGGCAGATAGCTGGGTATTCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
Db 828398 TCTCATGAGTCTGGCAGATAGCTGGGTATTCTTCTAGTAAACAGCTCGTCTTCTACTAG 828339
Qy 361 CAGATCTCAGACGTGGAGCTCAACGACGACGCGACCTAGCCCTCTCCACCACCGTT 420
Db 828338 CAGATCTCAGACGTGGAGCTCAACGACGACGCGACCTAGCCCTCTCCACCACCGTT 828279
Qy 421 TGATGATTATAGACTCAAGCGCAACACAGCTTACGATCTATCTTTACTCAACATCACT 480
Db 828278 TGATGATTATAGACTCAAGCGCAACACAGCTTACGATCTATCTTTACTCAACATCACT 828219
Qy 481 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATAATAAGGATAC 540
Db 828218 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATAATAAGGATAC 828159
Qy 541 AGCGCTACTGATAGGAACCGCAATCGCTCGGAGTGGGAACTTAAGATGCCGATGC 600
Db 828158 AGCGCTACTGATAGGAACCGCAATCGCTCGGAGTGGGAACTTAAGATGCCGATGC 828099
Qy 601 AGTTAAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTCGGATTAACCAAGCGAT 660
Db 828098 AGTTAAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTCGGATTAACCAAGCGAT 828039
Qy 661 TCTTGACTCTTTAGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
Db 828038 TCTTGACTCTTTAGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 827979
Qy 721 TGTAGCAACATAACAGAGCTGAGCTCTTAAAGAGATGCAAGATACCCAGTAGT 780
Db 827978 TGTAGCAACATAACAGAGCTGAGCTCTTAAAGAGATGCAAGATACCCAGTAGT 827919
Qy 781 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 840
Db 827918 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 827859
Qy 841 ACAGATAGAGAAGATGGAAATCGGATAGGATGATATTTTGCAGGACAGAAACGCTAG 900
Db 827858 ACAGATAGAGAAGATGGAAATCGGATAGGATGATATTTTGCAGGACAGAAACGCTAG 827799
Qy 901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACAGCTAAAGC 960
Db 827798 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACAGCTAAAGC 827739
Qy 961 AGCAATCGCTACTGCTTAAGACCAAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 1020
Db 827738 AGCAATCGCTACTGCTTAAGACCAAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 827679
Qy 1021 AATTCTTCAAGAACGGAACAAATGGTAATACAGCTGAGAAAGATCTTAAATAATCA 1080
Db 827678 AATTCTTCAAGAACGGAACAAATGGTAATACAGCTGAGAAAGATCTTAAATAATCA 827619
Qy 1081 ACCTGCAGATGTTCTGTGTTTCCAAATCCAGGAATACAGTTGGAGGCTCCAGCAACA 1140
Db 827618 ACCTGCAGATGTTCTGTGTTTCCAAATCCAGGAATACAGTTGGAGGCTCCAGCAACA 827559
Qy 1141 AGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAAAATGAGAC 1200
Db 827558 AGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTTTAGATGATGCTGAAAAATGAGAC 827499
Qy 1201 CGCTTCCATTTTGAATCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC 1260
Db 827498 CGCTTCCATTTTGAATCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC 827439
Qy 1261 TGATTTCTCAAGCTGCCCAACAGAGCTCGCAGCACAGCTAGAGCGAGGCGAAAGCGCTGG 1320

Db 827438 TGATTTCTCAAGCTGCCCAACAGGAGCTCGCAGCAAAAGCTAGACGCAAAAGCGCTGG 827379
Qy 1321 AGATGACAGTGTCTCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 1380
Db 827378 AGATGACAGTGTCTCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 827319
Qy 1381 TAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTG- 1439
Db 827318 TAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTG- 827259
Qy 1440 TTGCTAGCGCGAGGAGT 1455
Db 827258 TTGCTAGCGCGAGGAGT 827243

RESULT 5
US-09-841-132-385
; Sequence 385, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yaser A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469CB
; CURRENT APPLICATION NUMBER: US/09/841132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 509
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 385
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-385



Query Match 93.1%; Score 1355; DB 9; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGGTTAATCTTATGGTCCAGGTCCTATAGCAAGAAACAGAACGCACACCTCCGCGAGT 160
Db 1 ATGGTTAATCTTATGGTCCAGGTCCTATAGCAAGAAACAGAACGCACACCTCCGCGAGT 60
Qy 161 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATTAAGAGTGGGAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATTAAGAGTGGGAGCTCAAGAATA 120
Qy 221 GCAGTGGGAGAGCTTAAGCTTAAGAAATCTAAGACCGATTTCTGTAGAGCGATGGAGCATC 280
Db 121 GCAGTGGGAGAGCTTAAGCTTAAGAAATCTAAGACCGATTTCTGTAGAGCGATGGAGCATC 180
Qy 281 TTGGCTTCTGCAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGGCTTCTGCAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGT 240
Qy 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGATCAACGACAGCGACCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGATCAACGACAGCGACCGACCT 300
Qy 401 AGCGTCTCTCCACCCAGCTTTGATGATTATAGACTCAAGCGCAAAACAGCTTAGCATACT 460
Db 301 AGCGTCTCTCCACCCAGCTTTGATGATTATAGACTCAAGCGCAAAACAGCTTAGCATACT 360
Qy 461 ATCTTTACTCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCGCTCCAGGATGCT 520
Db 361 ATCTTTACTCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCGCTCCAGGATGCT 420
Qy 521 GTCACTAATAATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGG 580
Db 421 GTCACTAATAATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGG 480
Qy 581 GAACTAAGATGCCGATGCGAGTTAAGTTGGCGCGCAATTAAGATTTAGCGGAATAT 640

481 GAACTAAGAAATGCCGATGAGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 540
641 GCTTCGGATAACCAAGCGATCTTGTACTTTAGGTAAACTGACTTCTTCGACCTCTTA 700
541 GCTTCGGATAACCAAGCGATCTTGTACTTTAGGTAAACTGACTTCTTCGACCTCTTA 600
701 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 760
601 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 660
761 ATGCAAGATTAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCTCAATCTTTAGTTGAT 820
661 ATGCAAGATTAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCTCAATCTTTAGTTGAT 720
821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 880
721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 780
881 TTTGCGAGCAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 940
781 TTTGCGAGCAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 840
941 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
841 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
1001 AAAAAGTCCCGGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060
901 AAAAAGTCCCGGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 960
1061 AAGATCTTAAATAATCAACCTGCAAGTGGTCTGATGTTTCAATCCAGGAACCTACA 1120
961 AAGATCTTAAATAATCAACCTGCAAGTGGTCTGATGTTTCAATCCAGGAACCTACA 1020
1121 GTTGGAGGCTCCAGCAACCAAGAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTA 1180
1021 GTTGGAGGCTCCAGCAACCAAGAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTA 1080
1181 GATGATGCTGAAATGAGACCGGTTCCATTTGATGTTCTGGGTTTCTGAGATGATTCAC 1240
1081 GATGATGCTGAAATGAGACCGGTTCCATTTGATGTTCTGGGTTTCTGAGATGATTCAC 1140
1241 ATGTTCAATACGGAATATCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGACGACAGCT 1300
1141 ATGTTCAATACGGAATATCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGACGACAGCT 1200
1301 AGAGCAGGAAAGCCGCTGGAGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1360
1201 AGAGCAGGAAAGCCGCTGGAGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1260
1361 GCTTTAGAGCGGCTPAGGTAAAGCTGGGCAACAAACAGGCGATPCTCAATGTTTAGGA 1420
1261 GCTTTAGAGCGGCTPAGGTAAAGCTGGGCAACAAACAGGCGATPCTCAATGTTTAGGA 1320
1421 CAGATGCTTCTGCTGCTGTTGAGCGCAGGAGT 1455
1321 CAGATGCTTCTGCTGCTGTTGAGCGCAGGAGT 1355

RESULT 6
US-09-841-260-63
; Sequence 63, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; FILE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-63

Query Match 93.1%; Score 1355; DB 10; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGCTTTAATCTTATTCGTCAGGTCTCTATAGCAAAACAGAACGACACCTCCGCGAGAT 160
Db 1 ATGCTTTAATCTTATTCGTCAGGTCTCTATAGCAAAACAGAACGACACCTCCGCGAGAT 60
QY 161 CTTTCTGCTCAAGGATTGGAGGCGAGTGCAGCAAAATAGAGTGCAGGAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTGGAGGCGAGTGCAGCAAAATAGAGTGCAGGAGCTCAAGAATA 120
QY 221 GCAGGTGCGGAAAGCTTAAGCTTAAAGAAATCTAAGACCGAATCTGTAGAGCGATGAGCATC 280
Db 121 GCAGGTGCGGAAAGCTTAAGCTTAAAGAAATCTAAGACCGAATCTGTAGAGCGATGAGCATC 180
QY 281 TTGCGTTCTGCAAGTAAATGCTCTCATGACTCTGCGAGATAAGCTGGTATTGCTTCTAGT 340
Db 181 TTGCGTTCTGCAAGTAAATGCTCTCATGACTCTGCGAGATAAGCTGGTATTGCTTCTAGT 240
QY 341 AACAGTCTGCTCTTCTACTAGCAGATCTGCAGAGCTGGAATCAACGACGACGCGCACCT 400
Db 241 AACAGTCTGCTCTTCTACTAGCAGATCTGCAGAGCTGGAATCAACGACGACGCGCACCT 300
QY 401 AGCCTCTCTCCAGCCACGTTTGAATTAAGACTCAAGCGCAAAACAGCTTACGATAT 460
Db 301 AGCCTCTCTCCAGCCACGTTTGAATTAAGACTCAAGCGCAAAACAGCTTACGATAT 360
QY 461 ATCTTTACCTCAACATCACTAGTGCATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 520
Db 361 ATCTTTACCTCAACATCACTAGTGCATACAGGCTGCTTTGGTGGAGCTCCAGGATGCT 420
QY 521 GTCACATAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 580
Db 421 GTCACATAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 480
QY 581 GAAACTAAGATGCCCGATGCAAGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 640
Db 481 GAAACTAAGATGCCCGATGCAAGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 540
QY 641 GCTTCGGATAACCAAGCGGATCTTGTACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 700
Db 541 GCTTCGGATAACCAAGCGGATCTTGTACTCTTTAGGTAAACTGACTTCTTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 660
QY 761 ATGCAAGATAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTAGTCCAGGAGAAACCGCTGCAATCTCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 880
Db 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 780
QY 881 TTTGCGAGCAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 940
Db 781 TTTGCGAGCAGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 840
QY 941 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
QY 1001 AAAAAGTCCCGGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060

Db 901 AAAAAGTTCCCGACTCTCCAAATCTTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAG 960
Qy 1061 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACA 1120
Db 961 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACA 1020
Qy 1121 GTTGGAGCTCCAGCAACAAAGGAAGTAGTATGGTAGTATTCGTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGCTCCAGCAACAAAGGAAGTAGTATGGTAGTATTCGTGTTTCCATGCTGTTA 1080
Qy 1181 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1240
Db 1081 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATTCCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACCAAGCT 1300
Db 1141 ATGTTCAATACGGAATTCCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACCAAGCT 1200
Qy 1301 AGACAGGAAAGCGCTGGAGATGACAGTCTGTCGTCAGCGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGACAGGAAAGCGCTGGAGATGACAGTCTGTCGTCAGCGCTGGCAGATGCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAAGGCGCATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAAGGCGCATCTCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1455
Db 1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1355

RESULT 7
US-10-007-693-63
; Sequence 63, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.51-5C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-63

Query Match 93.1%; Score 1355; DB 13; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGGTTAATCTATTGGTCCAGGTCCTATAGACGAAACAGAACGCACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTATTGGTCCAGGTCCTATAGACGAAACAGAACGCACACCTCCCGCAGAT 60
Qy 161 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATAAGAGTGGCGAAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATAAGAGTGGCGAAGCTCAAGAATA 120
Qy 221 GCAGTGGGGAAGCTTAAGCTTAAGAACTTAAGCCGATTTGTAGAGCGATGGAGCATC 280
Db 121 GCAGTGGGGAAGCTTAAGCTTAAGAACTTAAGCCGATTTGTAGAGCGATGGAGCATC 180
Qy 281 TTGCGTTCTGAGTGAATGCTCTCATGATCTGGCAGATAGCTGGGTTATGCTTTAGT 340
Db 181 TTGCGTTCTGAGTGAATGCTCTCATGATCTGGCAGATAGCTGGGTTATGCTTTAGT 240
Qy 341 AACAGTCTGCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGACCGACCT 400
Db 241 AACAGTCTGCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGACCGACCT 300

Qy 401 ACGCCTCTCCACCACCAGTTTGTATGATTAAAGACTCAAGCGCAACAGCTTACGATCT 460
Db 301 ACGCCTCTCCACCACCAGTTTGTATGATTAAAGACTCAAGCGCAACAGCTTACGATCT 360
Qy 461 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
Db 361 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
Qy 521 GTCACCTAATAATAAAGGATACAGCGGCTACTGTAGAGGAAACCCCAATCGCTGGGAGTGG 580
Db 421 GTCACCTAATAATAAAGGATACAGCGGCTACTGTAGAGGAAACCCCAATCGCTGGGAGTGG 480
Qy 581 GAACTAAGAAATGCGGATGAGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 640
Db 481 GAACTAAGAAATGCGGATGAGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT 540
Qy 641 GCTTCGGATTAACCAAGGATTTCTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTA 700
Db 541 GCTTCGGATTAACCAAGGATTTCTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTA 600
Qy 701 CAGGCTCTCTTCTCCAACTCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 760
Db 601 CAGGCTCTCTTCTCCAACTCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAG 660
Qy 761 ATGCAAGATTAACCAAGTATGCCAGGGAACCGCTGCAATTCGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATTAACCAAGTATGCCAGGGAACCGCTGCAATTCGCTCAATCTTTAGTTGAT 720
Qy 821 CAGACAGATCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 880
Db 721 CAGACAGATCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGGATGCATAT 780
Qy 881 TTTGCGAGCAGAACCGCTAGCTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 940
Db 781 TTTGCGAGCAGAACCGCTAGCTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 840
Qy 941 AACATAGTTTCAAGTAAAGCAGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGTTTCAAGTAAAGCAGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAG 900
Qy 1001 AAAAAGTTCCCGACTCTCCAAATCTTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAG 1060
Db 901 AAAAAGTTCCCGACTCTCCAAATCTTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAG 960
Qy 1061 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACA 1120
Db 961 AAAGATCTTAAATAATCAAAACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACA 1020
Qy 1121 GTTGGAGCTCCAGCAACAAAGGAAGTAGTATGGTAGTATTCGTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGCTCCAGCAACAAAGGAAGTAGTATGGTAGTATTCGTGTTTCCATGCTGTTA 1080
Qy 1181 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1240
Db 1081 GATGATCTGAAATGAGACCGCTTCCATTTGATGTTCTGGTTTCGTCAGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATTCCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACCAAGCT 1300
Db 1141 ATGTTCAATACGGAATTCCTGATTTCTCAAGCTGCCCAACAGGAGCTCGCACCAAGCT 1200
Qy 1301 AGACAGGAAAGCGCTGGAGATGACAGTCTGTCGTCAGCGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGACAGGAAAGCGCTGGAGATGACAGTCTGTCGTCAGCGCTGGCAGATGCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAAGGCGCATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAAGGCGCATCTCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1455
Db 1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGT 1355

481	DB	GAAACATAGAAAGTCCGAGTGAAGTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATAT	541
641	QY	GCTTCGGATTAACCAAGCGATTCTTTGACTCTTTTAGGTAAACTGACTTCCTTCGACCTCTTA	700
541	DB	GCTTCGGATTAACCAAGCGATTCTTTGACTCTTTTAGGTAAACTGACTTCCTTCGACCTCTTA	600
701	QY	CAGGCTGCTCTTCTCCAACTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAGAG	760
601	DB	CAGGCTGCTCTTCTCCAACTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAGAG	660
761	QY	ATCAAGATTAACCCAGTAGTCCCAGGGAACGCCCTGCAATGTGCTCAATCTTTAGTTGAT	820
661	DB	ATCAAGATTAACCCAGTAGTCCCAGGGAACGCCCTGCAATGTGCTCAATCTTTAGTTGAT	720
821	QY	CAGACAGATGCTACAGCGCACACAGATAGAGAAAGATGGAATGGGATTAGGATGCAAT	880
721	DB	CAGACAGATGCTACAGCGCACACAGATAGAGAAAGATGGAATGGGATTAGGATGCAAT	780
881	QY	TTTTCGAGGACAGAAACGCTAGTGGAGCTGTAGAAAACTCTAAATCTAATAACAGTATAAGC	940
781	DB	TTTTCGAGGACAGAAACGCTAGTGGAGCTGTAGAAAACTCTAAATCTAATAACAGTATAAGC	840
941	QY	AACATAGATTACAGTAAGCAGCAATCGCTACTGCTAGACACAAATAGCTGAAGCTCAG	1000
841	DB	AACATAGATTACAGTAAGCAGCAATCGCTACTGCTAGACACAAATAGCTGAAGCTCAG	900
1001	QY	AAAAAGTTCGCCGACTCTCCAAATCTTTCAGAGAGCGGAACAAATGGTAAATACAGGCTGAG	1060
901	DB	AAAAAGTTCGCCGACTCTCCAAATCTTTCAGAGAGCGGAACAAATGGTAAATACAGGCTGAG	960
1061	QY	AAGATGCTTTAAATAATCAACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACTACA	1120
961	DB	AAGATGCTTTAAATAATCAACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAACTACA	1020
1121	QY	GTTGGAGGCTCCAAAGCAACAAGGAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTCTTA	1180
1021	DB	GTTGGAGGCTCCAAAGCAACAAGGAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTCTTA	1080
1181	QY	GATGATGCTGAAATAGACACCGCTTCCATTTTGATGCTGGGTTTCGTAGATGATTAC	1240
1081	DB	GATGATGCTGAAATAGACACCGCTTCCATTTTGATGCTGGGTTTCGTAGATGATTAC	1140
1241	QY	ATGTTCAATACGGAATAATCCTGATTTCTCAAGTGCCTCAACACAGGAGCTCGCAGACAAAGCT	1300
1141	DB	ATGTTCAATACGGAATAATCCTGATTTCTCAAGTGCCTCAACACAGGAGCTCGCAGACAAAGCT	1200
1301	QY	AGACGAGGAAAGCCGCTGGAGATGACAGTGCTGCTGACGCGTGGCAGATGCTCAGAAA	1360
1201	DB	AGACGAGGAAAGCCGCTGGAGATGACAGTGCTGCTGACGCGTGGCAGATGCTCAGAAA	1260
1361	QY	GCPTTAGAAGCGCTCTAGTTAAAGCTGGGCAACAACAGGGCATCTCAATGCTTTAGGA	1420
1261	DB	GCPTTAGAAGCGCTCTAGTTAAAGCTGGGCAACAACAGGGCATCTCAATGCTTTAGGA	1320
1421	QY	CAGATCGCTTCGCTGCTGTTGTGAGCGCAGGAGT	1455
1321	DB	CAGATCGCTTCGCTGCTGTTGTGAGCGCAGGAGT	1355

```
RESULT 9
US-09-564
; Sequence
; Publications
; GENERAL
; APPLIC
; APPLIC
; APPLIC
; APPLIC
; TITLE
; TITLE
; FILE
; CURRENT
; CURRENT
```

RESULT 8					
US-10-312-273-14					
; Sequence 14, Application US/10312273					
; Publication No. US20040005667A1					
; GENERAL INFORMATION:					
; APPLICANT: CHIRON SpA					
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE					
; FILE REFERENCE: P025035WO					
; CURRENT APPLICATION NUMBER: US/10/312.273					
; CURRENT FILING DATE: 2002-12-20					
; PRIOR APPLICATION NUMBER: 0016363.4					
; PRIOR FILING DATE: 2000-07-03					
; PRIOR APPLICATION NUMBER: 0017047.2					
; PRIOR FILING DATE: 2000-07-11					
; PRIOR APPLICATION NUMBER: 0017993.8					
; PRIOR FILING DATE: 2000-07-21					
; PRIOR APPLICATION NUMBER: 0019368.0					
; PRIOR FILING DATE: 2000-08-07					
; PRIOR APPLICATION NUMBER: 0020440.4					
; PRIOR FILING DATE: 2000-08-18					
; PRIOR APPLICATION NUMBER: 0022583.9					
; PRIOR FILING DATE: 2000-09-14					
; PRIOR APPLICATION NUMBER: 0027549.5					
; PRIOR FILING DATE: 2000-11-10					
; PRIOR APPLICATION NUMBER: 0031706.5					
; PRIOR FILING DATE: 2000-12-22					
; NUMBER OF SEQ ID NOS: 664					
; SOFTWARE: SeqWin99, version 1.02					
; SEQ ID NO 14					
; LENGTH: 1956					
; TYPE: DNA					
; ORGANISM: Chlamydia pneumoniae					
US-10-312-273-14					
Query Match 93.1%; Score 1355; DB 15; Length 1956;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	101	ATGGTTAAATCCTATTGTCAGTCCAGTCTCTATAGACGAAAACAGAACGCACCTCCCGCAGAT	160		
Db	1	ATGGTTAAATCCTATTGTCAGTCCAGTCTCTATAGACGAAAACAGAACGCACCTCCCGCAGAT	60		
QY	161	CTTTCTCTCAAGGATTGGAGCGAGTGCGACAATAAGAGTGGGAAGCTCAAAGAATA	220		
Db	61	CTTTCTCTCAAGATTGGAGCGAGTGCGACAATAAGAGTGGGAAGCTCAAAGAATA	120		
QY	221	GCGAGTGGGAAGCTTAAGCCTTAAAGATCTTAAGACCGATCTGTAGAGCGATGGAGCATC	280		
Db	121	GCGAGTGGGAAGCTTAAGCCTTAAAGATCTTAAGACCGATCTGTAGAGCGATGGAGCATC	180		
QY	281	TTCGGTTCTGCAGTGATGCTCTCATAGCTGGCAGATAGCTGGGTATGCTTCTTAGT	340		
Db	181	TTCGGTTCTGCAGTGATGCTCTCATAGCTGGCAGATAGCTGGGTATGCTTCTTAGT	240		
QY	341	AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACCT	400		
Db	241	AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACCT	300		
QY	401	ACGGCTCTCCACCCACGTTTGATGATTAAAGCTCAAGCGCAAAACAGCTTACGATACT	460		
Db	301	ACGGCTCTCTCCACCCACGTTTGATGATTAAAGCTCAAGCGCAAAACAGCTTACGATACT	360		
QY	461	ATCTTTTACCCTCAACATCACTAGCTGCATACAGGCTGCTTTGGTGACCTCCAGGATGCT	520		
Db	361	ATCTTTTACCCTCAACATCACTAGCTGCATACAGGCTGCTTTGGTGACCTCCAGGATGCT	420		
QY	521	GTCACTAATAATAAGGATACAGCGCTTACTGATGAGGAAAAACCGCAATCGCTGCCGAGTGG	580		
Db	421	GTCACTAATAATAAGGATACAGCGCTTACTGATGAGGAAAAACCGCAATCGCTGCCGAGTGG	480		
QY	581	GAACCTAAGAAATGCCGATGCAAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATAT	640		
Db					

```
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1749)
US-09-564-479-3

Query Match
 79.1%; Score 1151; DB 10; Length 1852;
Best Local Similarity 100.0%; Pred. No. 3.2e-289;
Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTCTACTAGCAGA 364
Db 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTCTACTAGCAGA 60
QY 365 TCTGCAGAGCTGGACTCAACGACAGCGACCGACCTAGCCCTCTCCACCCACGTTTGAT 424
Db 61 TCTGCAGAGCTGGACTCAACGACAGCGACCGACCTAGCCCTCTCCACCCACGTTTGAT 120
QY 425 GATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACTAGCT 484
Db 121 GATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACTAGCT 180
QY 485 GACATACAGGCTGCTTGGTGAAGCTCCAGGATGCTGTCACTAATAATAAGGATACAGG 544
Db 181 GACATACAGGCTGCTTGGTGAAGCTCCAGGATGCTGTCACTAATAATAAGGATACAGG 240
QY 545 GCTACTGATGAGGAACCGCAATCGCTGGAGTGGGAACTAAGAAATGCCGATCGAGTT 604
Db 241 GCTACTGATGAGGAACCGCAATCGCTGGAGTGGGAACTAAGAAATGCCGATCGAGTT 300
QY 605 AAAGTTGGCGCAAAATTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 664
Db 301 AAAGTTGGCGCAAAATTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATTCTT 360
QY 665 GACTCTTTAGTAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCCCAATCTGTA 724
Db 361 GACTCTTTAGTAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCCCAATCTGTA 420
QY 725 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCA 784
Db 421 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCA 480
QY 785 GGGAAACCCCTGGAAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAG 844
Db 481 GGGAAACCCCTGGAAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAG 540
QY 845 ATAGAGAAGATGGAAATCGGATAGGAGTGCATATTTTCAGGACAGAACTAGTGA 904
Db 541 ATAGAGAAGATGGAAATCGGATAGGAGTGCATATTTTCAGGACAGAACTAGTGA 600
QY 905 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGCAGCA 964
Db 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAGCAGCA 660
QY 965 ATCGCTACTGCTAAGCACAATAATAGCTGAAGCTCAGAAAAGTCCCGGACTCTCCAAT 1024
Db 661 ATCGCTACTGCTAAGCACAATAATAGCTGAAGCTCAGAAAAGTCCCGGACTCTCCAAT 720
QY 1025 CTTCAAGAGCGGAACAAATGTTAATACAGGCTGAGAAAGATCTTAAATAATCAAACT 1084
Db 721 CTTCAAGAGCGGAACAAATGTTAATACAGGCTGAGAAAGATCTTAAATAATCAAACT 780
QY 1085 GCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAGGA 1144
Db 1085 GCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAGGA 1144

Db 781 GCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAGGA 840
QY 1145 AGTAGTATTGGTAGTATTTCGTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 1204
Db 841 AGTAGTATTGGTAGTATTTCGTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 900
QY 1205 TCCATTTTGTGCTGCTGGTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 1264
Db 901 TCCATTTTGTGCTGCTGGTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 960
QY 1265 TCTCAAGCTGCCAAACAGGAGCTCGCAGCACAAGCTAGACAGCGAAGCGCTGGAGAT 1324
Db 961 TCTCAAGCTGCCAAACAGGAGCTCGCAGCACAAGCTAGACAGCGAAGCGCTGGAGAT 1020
QY 1325 GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAGCGGCTCTAGGTAAA 1384
Db 1021 GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
QY 1385 GCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1444
Db 1081 GCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
QY 1445 AGCGCAGGAGT 1455
Db 1141 AGCGCAGGAGT 1151

RESULT 10
US-09-905-119-1
; Sequence 1, Application US/09905119
; Patent No. US2002091096A1
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew D
; APPLICANT: Oomen, Raymond P
; APPLICANT: Dunn, Pamela L
; TITLE OF INVENTION: Chlamydia Antigens and Corresponding DNA Fragments and
; FILE REFERENCE: 19721-010 DIV
; CURRENT APPLICATION NUMBER: US/09/905,119
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/106,037
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/154,658
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/427,501
; PRIOR FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1366)
US-09-905-119-1

Query Match
 11.7%; Score 170; DB 9; Length 1550;
Best Local Similarity 100.0%; Pred. No. 2e-33;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACACAGGCTCGCANTTAATTATTAGTGAGAGCTTTTTTTTATTTTTT 60
Db 1381 ATAAATCTTTAAACACAGGCTCGCANTTAATTATTAGTGAGAGCTTTTTTTTATTTTTT 1440
QY 61 ATAATAAACTAAAGATTTTTTATTTTTTTAGTTTTTATGTTTAAATCCTATTGTGTC 120
Db 1441 ATAATAAACTAAAGATTTTTTATTTTTTTAGTTTTTATGTTTAAATCCTATTGTGTC 1500
QY 121 AGGTCTTATAGACGAACAGAACGACACCTCCCGCAGATCTTCTGCTC 170
Db 1501 AGGTCTTATAGACGAACAGAACGACACCTCCCGCAGATCTTCTGCTC 1550
```


Search completed: March 23, 2004, 19:25:22
Job time : 509.565 secs

10-608-559-5.rnpb

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:15 ; Search time 5658.52 Seconds
(without alignments)
11152.641 Million cell updates/sec

Title: US-10-608-559-5
Perfect score: 1456
Sequence: 1 aataatctttaaaacagg.....ctgttgagcaggaagta 1456

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_man:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1456	100.0	1456	6	AX045135 Sequence
2	1456	100.0	2238	6	AX045137 Sequence
3	1455	99.9	2156	6	AX045131 Sequence
4	1455	99.9	10421	1	AE001654 Chlamydia
5	1455	99.9	12548	1	AE002165 Chlamydia
6	1455	99.9	300566	1	AE017159 Chlamydia
7	1455	99.9	300550	1	AP002547 Chlamydia
8	1444	99.2	2148	1	CHT76KDA
9	1444	99.2	110000	6	AR310754_08
10	1355	93.1	1956	6	AX338293 Sequence
11	1355	93.1	1956	6	AX349491 Sequence
12	1355	93.1	1956	6	AX361992 Sequence
13	1151	79.1	1852	6	AX045133 Sequence
14	204.2	14.0	304769	1	AE016994 Chlamydia
15	170	11.7	1550	6	BD243971 Chlamydia
16	170	11.7	1550	6	AR212731 Sequence
17	170	11.7	1550	6	AR428600 Sequence
18	148.4	10.2	150	6	AR144067 Sequence
19	55.4	3.8	2000	6	AX655393 Sequence
20	54.4	3.7	1941	6	AX362019 Sequence
21	54.4	3.7	1944	6	AX786284 Sequence
22	54.4	3.7	7218	6	I66494 Sequence 14
23	54.4	3.7	10354	1	AE001333 Chlamydia
24	51.6	3.5	2205	6	AR317736 Sequence
25	51.4	3.5	1537	6	BD264444 Compounds
26	51.4	3.5	1537	6	AR229257 Sequence
27	51.4	3.5	1537	6	AR308168 Sequence
28	51.4	3.5	1537	6	AR321875 Sequence
29	51.4	3.5	1537	6	AX155814 Sequence
30	51.4	3.5	1537	6	AX361664 Sequence
31	51.4	3.5	154160	2	BX571709 Danio rerio
32	50.6	3.5	423	1	MC049
33	50.2	3.4	14698	1	AE002357 Chlamydia
34	49.4	3.4	5098	1	MCU51235 Mycoplasma
35	49	3.4	1990	6	AR268644 Sequence
36	49	3.4	2085	6	AR008322 Sequence
37	49	3.4	2085	6	AR027306 Sequence
38	49	3.4	2085	6	AR035771 Sequence
39	49	3.4	2085	6	AR037631 Sequence
40	49	3.4	2085	6	AR078816 Sequence
41	49	3.4	2085	6	AR079063 Sequence
42	49	3.4	2085	6	AR084296 Sequence
43	49	3.4	2085	6	I16740 Sequence 1
44	49	3.4	2085	6	I70472 Sequence 1
45	49	3.4	2085	6	I92332 Sequence 1

ALIGNMENTS

RESULT 1	AX045135	Sequence 5 from Patent WO0066739.	1456 bp	DNA	linear	PAT 24-NOV-2000
LOCUS	AX045135	Sequence 5 from Patent WO0066739.				
DEFINITION	AX045135					
ACCESSION	AX045135					
VERSION	AX045135.1	GI:11343734				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						

```
FEATURES
source
    Aventis Pasteur Limited (CA)
    Location/Qualifiers
        1..1456
            /organism="Chlamydomophila pneumoniae"
            /mol_type="unassigned DNA"
            /db_xref="taxon:83558"
            101..>1452
                /note="unnamed protein product"
                /codon_start=1
                /transl_table=11
                /protein_id="CAC17379.1"
                /db_xref="GI:11343735"
                /db_xref="EMBL:CA17379"
                /translation="MNPFGPGFIDETERTPPADLSAQGLEASAAANKSAEARIAGAE
                AKPKESKDSVERKILSSAVNALMSLADKLGIASSNSNSSSTSRSDVSTTATATPT
                PPTFDYKQATYDTFTSTSLADIQAALVSLQDAVTNIKDTAAATDEETAIAAEW
                ETKNADAVKVGACIITELAKYASDNOAILDSKLKLSFDLLQALQLSVANNKAAELL
                KEMONPVPGKTPALQSLVDQTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSN
                NSISNIDSAKALIAKQIAEAKKFPDPSILQAEQVIOREKDLNKRFPADGSDV
                PNFGTIVGSGKQSGSSIGSIRVSMLLDDAENETASILMSGFQMIHMENTNPDSQAA
                QQELAAQARAARAAAGSDSAAALADAQKALEAALGKAGQQQQLINALQGIASAAVYSA
                G"
ORIGIN
    Query Match      100.0%; Score 1456; DB 6; Length 1456;
    Best Local Similarity 100.0%; Pred. No. 2.4e-301;
    Matches 1456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTTTTTATTTTTT 60
    DB 1 ATAAATCTTTAAACAGCGCTCGCATTAATATTAGTCAGAGCTTTTTTTTATTTTTT 60
    QY 61 ATATAAACTAAAGATTTTATTTATTTTGTAGTTTTTATGGTAAATCCATTGGTCC 120
    DB 61 ATATAAACTAAAGATTTTATTTATTTTGTAGTTTTTATGGTAAATCCATTGGTCC 120
    QY 121 AGTCTCTATAGAGCAACAGCGACACCTCCCGAGATCTTCTGCTCAGAGGATTGGA 180
    DB 121 AGTCTCTATAGAGCAACAGCGACACCTCCCGAGATCTTCTGCTCAGAGGATTGGA 180
    QY 181 GCGGAGTCAGCAAAATAAGAGTCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 240
    DB 181 GCGGAGTCAGCAAAATAAGAGTCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 240
    QY 241 TAAAGAACTAAGACCGATTCTGTAGAGCGATGGAGCATCTTCGCGTTCTGCAGTGAATGC 300
    DB 241 TAAAGAACTAAGACCGATTCTGTAGAGCGATGGAGCATCTTCGCGTTCTGCAGTGAATGC 300
    QY 301 TCTCATGAGTCGCGAGATAAGCTGGTATTCGTTCTAGTAAACAGCTCGTCTTCTACTAG 360
    DB 301 TCTCATGAGTCGCGAGATAAGCTGGTATTCGTTCTAGTAAACAGCTCGTCTTCTACTAG 360
    QY 361 CAGATCTGCAGACGTGGATCAACGACAGCGACCGACCTACGCCCTCTCCACCCACGTT 420
    DB 361 CAGATCTGCAGACGTGGATCAACGACAGCGACCGACCTACGCCCTCTCCACCCACGTT 420
    QY 421 TGATGATTAAGACTCAGCGCAACACGCTTACGATCTATCTTTACCTCAACATCACT 480
    DB 421 TGATGATTAAGACTCAGCGCAACACGCTTACGATCTATCTTTACCTCAACATCACT 480
    QY 481 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACCTAATATAAGGATAC 540
    DB 481 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGTCACCTAATATAAGGATAC 540
    QY 541 AGCGGCTACTGATAGGAAACCGCAATCGTCGGAGTGGAAACTAAGAATGCCGATGC 600
    DB 541 AGCGGCTACTGATAGGAAACCGCAATCGTCGGAGTGGAAACTAAGAATGCCGATGC 600
    QY 601 AGTTAAAGTTGGCGCCAAATACAGAAATAGCGAAATATGCTTCGGATAACACAGCGAT 660
    DB 601 AGTTAAAGTTGGCGCCAAATACAGAAATAGCGAAATATGCTTCGGATAACACAGCGAT 660
    QY 661 TCTTGACTCTTTAGGTAAACTGACTCTTCCTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
    DB 661 TCTTGACTCTTTAGGTAAACTGACTCTTCCTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
    QY 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
    DB 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
    QY 781 CCCAGGAAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 840
    DB 781 CCCAGGAAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCGAC 840
    QY 841 ACAGATAGAGAAAGATGGAATCGGATTAGGATGCAATTTTGCAGACAGACGCTAG 900
    DB 841 ACAGATAGAGAAAGATGGAATCGGATTAGGATGCAATTTTGCAGACAGACGCTAG 900
    QY 901 TGGAGCTCTAGAAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTTCAGCTAAAGC 960
    DB 901 TGGAGCTCTAGAAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTTCAGCTAAAGC 960
    QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGCAGCTCTCC 1020
    DB 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGCAGCTCTCC 1020
    QY 1021 AATTCTTCAAGACGGGAACAAATGGTAAATACAGCTGAGAAAGATCTTAAATAATATCAA 1080
    DB 1021 AATTCTTCAAGACGGGAACAAATGGTAAATACAGCTGAGAAAGATCTTAAATAATATCAA 1080
    QY 1081 ACCTGCAGATGGTTCGATGTTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
    DB 1081 ACCTGCAGATGGTTCGATGTTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
    QY 1141 AGGAAGTAGTATGTTGTTAGTATTCGTTTCCATGCTGTAGATGCTGTAATGAGAC 1200
    DB 1141 AGGAAGTAGTATGTTGTTAGTATTCGTTTCCATGCTGTAGATGCTGTAATGAGAC 1200
    QY 1201 CGTTCCTCAATTTTGTGCTGGGTTTCGTCAGATGATTCACATGTTTCAATACGGAATAATCC 1260
    DB 1201 CGTTCCTCAATTTTGTGCTGGGTTTCGTCAGATGATTCACATGTTTCAATACGGAATAATCC 1260
    QY 1261 TGATTTCTCAAGCTGCCCAACAGGAGCTTCGACAGCAAGCTAGACGAGCAAGCGCTGG 1320
    DB 1261 TGATTTCTCAAGCTGCCCAACAGGAGCTTCGACAGCAAGCTAGACGAGCAAGCGCTGG 1320
    QY 1321 AGATGACAGTCTGCTGCGAGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
    DB 1321 AGATGACAGTCTGCTGCGAGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
    QY 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGGCTTCTGCTGCTGT 1440
    DB 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGGCTTCTGCTGCTGT 1440
    QY 1441 TGTGAGCGCAGAGTA 1456
    DB 1441 TGTGAGCGCAGAGTA 1456
RESULT 2
AX045137
LOCUS
DEFINITION
    Sequence 7 from Patent WO0066739.
ACCESSION
    AX045137
VERSION
    AX045137.1
KEYWORDS
    GI:11343735
SOURCE
    Chlamydomophila pneumoniae
    Chlamydomophila pneumoniae
    Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE
    1
AUTHORS
    Mordin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
TITLE
    i(chlamydia) antigens and corresponding dna fragments and uses thereof
JOURNAL
    Patent: WO 0066739-A 7 09-NOV-2000;
    Aventis Pasteur Limited (CA)
FEATURES
    Location/Qualifiers
```

source	1. .2238	
	/organism="Chlamydomophila pneumoniae"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:83558"	
CDS	766. .2238	
	/note="unnamed protein product"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="CAC17330.1"	
	/db_xref="GI:11343737"	
	/db_xref="REMTREMBL: CAC17330"	
	/translation="MWNPIGPGIDERTPPADLSAOGLEASANKSAAEORIAEAE APKESKTSVERMSILRAVNALMSLADKLIATSNSSSRSDVSTATATPT PPTFDYKTQAGTADITFTSLADIQAALVSDAVTNIKDNTAETETALAEW ETKNADVKGQITELAKYADNQLDLSGLKLTSLDQLAALQSVANNAEALL KEMQNPVPGKTPATQSLVDQDATQIEKDNARIDAYFAGQSNAGVENKSDV NSTSNIDSAKATATKATQIAEAKQKFPDPSPILOEAKQOMVIOAEKDKLNKPKADGSDV PNPPTVGSKQSGSISIRVSMILDDAENETASILMSGFRQIMHMENTENPDQAA QOELAAQAAKAAAGDSDAALADAKALEALKGAGQOQQLNLGQIASAAVVA GVLPLQOVWIRARYQAYVEQKLISEDLSNVAVDHHHHH"	
ORIGIN		
Query Match	100.0%; Score 1456; DB 6; Length 2238;	
Best Local Similarity	100.0%; Pred. No. 2.3e-301;	
Matches 1456; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTTT	60
DB	666 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTTT	725
QY	61 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTTT	120
DB	726 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTTT	785
QY	121 AGTCTCTATAGACGAACAGAGCGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	180
DB	786 AGTCTCTATAGACGAACAGAGCGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA	845
QY	181 GCGAGTGCAGCAATAGAGTCCGAAGCTCAAGAAATAGAGGTGCGGAAGCTAAGCC	240
DB	846 GCGAGTGCAGCAATAGAGTCCGAAGCTCAAGAAATAGAGGTGCGGAAGCTAAGCC	905
QY	241 TAAAGATCTTAAGCCGATCTCTAGAGGATGAGCATCTTCTGCTTCTGAGTGAATGC	300
DB	906 TAAAGATCTTAAGCCGATCTCTAGAGGATGAGCATCTTCTGCTTCTGAGTGAATGC	965
QY	301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG	360
DB	966 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG	1025
QY	361 CAGATCTGAGAGTGGACTCAAGCAGGACCGACCTACGCTCTCCACCCAGCTT	420
DB	1026 CAGATCTGAGAGTGGACTCAAGCAGGACCGACCTACGCTCTCCACCCAGCTT	1085
QY	421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT	480
DB	1086 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT	1145
QY	481 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGCTCACTAATATAAGGATAC	540
DB	1146 AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGCTCACTAATATAAGGATAC	1205
QY	541 AGCGGCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAACCTAAGNATGCCATGC	600
DB	1206 AGCGGCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAACCTAAGNATGCCATGC	1265
QY	601 AGTTAAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT	660
DB	1266 AGTTAAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGAT	1325
QY	661 TCTTGACTCTTTAGGTAAACTGACTTCTTCTGACCTCTTACAGGCTGCTCTTCTCAATC	720
DB	1326 TCTTGACTCTTTAGGTAAACTGACTTCTTCTGACCTCTTACAGGCTGCTCTTCTCAATC	1385

QY	721 TCTAGCAAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGT	780
DB	1386 TCTAGCAAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGT	1445
QY	781 CCCAGGGAACCGCTGCAATTCCTCAATTCCTTATGATCAGACAGATGCTACAGCGAC	840
DB	1446 CCCAGGGAACCGCTGCAATTCCTCAATTCCTTATGATCAGACAGATGCTACAGCGAC	1505
QY	841 ACAGATAGAGAAAGATGGAATTCGATAGGATGCATATTTTGGAGGACAGAAACGCTAG	900
DB	1506 ACAGATAGAGAAAGATGGAATTCGATAGGATGCATATTTTGGAGGACAGAAACGCTAG	1565
QY	901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGTAAAGC	960
DB	1566 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGTAAAGC	1625
QY	961 AGCAATCGTACTGCTTAAGACACAATAGCTCAAGCTCAGAAAAGTTCCCGGACTCTCC	1020
DB	1626 AGCAATCGTACTGCTTAAGACACAATAGCTCAAGCTCAGAAAAGTTCCCGGACTCTCC	1685
QY	1021 AATTCTTCAAGAGCGGAACAAATGGTATATACAGGCTGAGAAAGATCTTAAAAATATCAA	1080
DB	1686 AATTCTTCAAGAGCGGAACAAATGGTATATACAGGCTGAGAAAGATCTTAAAAATATCAA	1745
QY	1081 ACCTGAGATGCTTCTGATGTTCCAAATCCAGGAACACTAGTTGGAGGCTCCAAAGCAACA	1140
DB	1746 ACCTGAGATGCTTCTGATGTTCCAAATCCAGGAACACTAGTTGGAGGCTCCAAAGCAACA	1805
QY	1141 AGGAAGTACTATTGGTAGTATTCGTTTCATGCTGTAGATGATGCTGAAAATCAGAC	1200
DB	1806 AGGAAGTACTATTGGTAGTATTCGTTTCATGCTGTAGATGATGCTGAAAATCAGAC	1865
QY	1201 CGCTTCCATTTTATGCTGGGTTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC	1260
DB	1866 CGCTTCCATTTTATGCTGGGTTTTCGTCAGATGATTCATGTTCAATACGGAATAATCC	1925
QY	1261 TGATTTCTCAAGTCCCAACAGAGCTCCAGACACAGCTAGAGCGCAAGGCGGCTCG	1320
DB	1926 TGATTTCTCAAGTCCCAACAGAGCTCCAGACACAGCTAGAGCGCAAGGCGGCTCG	1985
QY	1321 AGATGACAGTGTCTGTCAGGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG	1380
DB	1986 AGATGACAGTGTCTGTCAGGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG	2045
QY	1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATCTTTAGGACAGATGCTTTCTGCTGCT	1440
DB	2046 TAAAGCTGGGCAACAAACAGGCGATCTCAATCTTTAGGACAGATGCTTTCTGCTGCT	2105
QY	1441 TGTGAGCGCAGGAGTA 1456	
DB	2106 TGTGAGCGCAGGAGTA 2121	

RESULT 3					
AX045131	Sequence 1 from Patent WO0066739.	2156 bp	DNA	linear	PAT 24-NOV-2000
LOCUS					
DEFINITION					
AX045131					
ACCESSION					
AX045131.1	GI:11343730				
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
Chlamydomophila pneumoniae					
Chlamydomophila pneumoniae					
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.					
1					
Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.					
i (Chlamydia) antigens and corresponding dna fragments and uses					
thereof					
Patent: WO 0066739-A 1 09-NOV-2000;					
Avantis Pasteur Limited (CA)					
Location/Qualifiers					
1. .2156					
/organism="Chlamydomophila pneumoniae"					

Handwritten signature and notes.

```
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
101..2056
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CACI7329.1"
/db_xref="GI:11343731"
/db_xref="REMTREMBL:CACI7329"
/translation="MVNPIGPGPIDETERTPPADLSAOGLEASANKSABEORLAGAE
AKPKESKDSVERWSILRSAYNALMSLADKLGIIASSNSSSTSRSDVSTATAPTP
PPPPDDYKTOQATYDPIFTSLADIQALVSLQDAYVNIKDTAATDEETAIAEAW
ETKNADAVVQKQATFELAKYSDNQALDLSLGLKTSFDLQALLOSVAANNKKAELL
KEMQNDPVPKPTPAIQSLVDQTDATQIEKQGNARIDAYFAGQNAKSGAVENAKSN
NSINISDAKAAIAFAKQIAEAKQKPPSPILQEAQOMVTOAEKOLNKPADGSDV
PNPITVGGSKQOQSSIGSIRVSMLLDAENETASILMSGFRQMHMENTNPDSQAA
QOELAAQARAAGADGAAALADAOKALEAALGKAGQOQILNALGQIASAAVUSA
GVPPAAASSIGSSVKQLYKTSKSGSDYKTOISAGYDAYKSDINAYGPARNDATDVI
NNVSTPALTRSPVPRARTARGPEKTDQALARVIGNSRTLGDIVYQVSALQSVNQIIQ
SNPQANNEEIRKQKTSVTKPPFGYPYVQLSNDSTQKFIKLESFLFAEGSRATAEIK
ALSPFTNSLFTQQVLVNIQSLYSGLQ"
```

ORIGIN

```
Query Match          99.9%; Score 1455; DB 6; Length 2156;
Best Local Similarity 100.0%; Pred. No. 3.8e-301; Indels 0; Gaps 0;
Matches 1455; Conservative 0; Mismatches 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTT 60
DB 1 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTT 60

QY 61 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTT 120
DB 61 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTTT 120

QY 121 AGGTCTATATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 180
DB 121 AGGTCTATATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 180

QY 181 GCGAGTGCGAGCAATATAGAGTGGGAGAGTCAAGAGATAGAGTGGGAGAGTCAAGC 240
DB 181 GCGAGTGCGAGCAATATAGAGTGGGAGAGTCAAGAGATAGAGTGGGAGAGTCAAGC 240

QY 241 TAAAGATCTAAGACCGATCTGTAGAGGATGGAGCATCTTCGCTTCTGAGTGAATGC 300
DB 241 TAAAGATCTAAGACCGATCTGTAGAGGATGGAGCATCTTCGCTTCTGAGTGAATGC 300

QY 301 TCTCATGAGTCTGCGAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 360
DB 301 TCTCATGAGTCTGCGAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAG 360

QY 361 CAGATCTGAGACGTGGATCTAACACAGCGACCGACCTAGCGCTCTCCACCCAGCTT 420
DB 361 CAGATCTGAGACGTGGATCTAACACAGCGACCGACCTAGCGCTCTCCACCCAGCTT 420

QY 421 TCGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT 480
DB 421 TCGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTCTTACCTCAACATCACT 480

QY 481 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACCTAATATAAGGATAC 540
DB 481 AGCTGACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACCTAATATAAGGATAC 540

QY 541 ACGGGCTACTGATAGGAAACCGCAATCTCGTGGAGTGGGAACTAAGATGCGGATGC 600
DB 541 ACGGGCTACTGATAGGAAACCGCAATCTCGTGGAGTGGGAACTAAGATGCGGATGC 600

QY 601 AGTTAAAGTTGGCGCAATATACAGATTAAGCAATATGCTTCGGATAACCAAGCGAT 660
DB 601 AGTTAAAGTTGGCGCAATATACAGATTAAGCAATATGCTTCGGATAACCAAGCGAT 660

QY 661 TCTTGACTCTTTAGGTAAACTGACTTCTTGTGACCTTTTACAGGCTGCTTCTTCTCAATC 720
DB 661 TCTTGACTCTTTAGGTAAACTGACTTCTTGTGACCTTTTACAGGCTGCTTCTTCTCAATC 720
```

```
Db 661 TCTTGACTCTTTAGGTAAACTGACTTCTTGTGACCTTTTACAGGCTGCTTCTTCTCAATC 720
QY 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGT 780
QY 781 CCCAGGGAAACCGCTCGCAATTTGCTCAATCTTTAGTTGATGATGATGATGATGATGATGATG 840
Db 781 CCCAGGGAAACCGCTCGCAATTTGCTCAATCTTTAGTTGATGATGATGATGATGATGATGATG 840
QY 841 ACAGATAGAGAAAGATGAAATGCGATTAGGATGCAATATTTTCAGGACAGACAGCGCTAG 900
Db 841 ACAGATAGAGAAAGATGAAATGCGATTAGGATGCAATATTTTCAGGACAGACAGCGCTAG 900
QY 901 TGGAGCTGTAGAAAATCTAAATCTAATAACAGTATAAGCAACATAGATTGATGATGATGATG 960
Db 901 TGGAGCTGTAGAAAATCTAAATCTAATAACAGTATAAGCAACATAGATTGATGATGATGATG 960
QY 961 AGCAATCGCTACTCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
Db 961 AGCAATCGCTACTCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
QY 1021 AATTCTTCAAGAGCGCAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 1021 AATTCTTCAAGAGCGCAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
QY 1081 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
Db 1081 ACCTGCGAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
QY 1141 AGGAAGTAGTATTGCTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC 1200
Db 1141 AGGAAGTAGTATTGCTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC 1200
QY 1201 CGCTTCCATTTTGTGTTCTGGTTTCGTCAGATGATTCATATGATGATGATGATGATGATGATG 1260
Db 1201 CGCTTCCATTTTGTGTTCTGGTTTCGTCAGATGATTCATATGATGATGATGATGATGATGATG 1260
QY 1261 TGATTTCAAGCTCCCAACAGGAGCTCGAGCAAGCTAGAGCAGGAGGAGGAGGAGGAGGAGGAGG 1320
Db 1261 TGATTTCAAGCTCCCAACAGGAGCTCGAGCAAGCTAGAGCAGGAGGAGGAGGAGGAGGAGGAGG 1320
QY 1321 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 1380
Db 1321 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 1380
QY 1381 TAAAGCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db 1381 TAAAGCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
QY 1441 TGTGAGCGCAGGAGT 1455
Db 1441 TGTGAGCGCAGGAGT 1455
```

RESULT 4

```
AE001654/c
LOCUS AE001654 10421 bp DNA linear BCT 01-DEC-2000
DEFINITION Chlamydia pneumoniae section 70 of 103 of the complete genome.
ACCESSION AE001654 AE001363
VERSION AE001654.1 GI:4377031
KEYWORDS
SOURCE Chlamydia pneumoniae CWL029
ORGANISM Chlamydia pneumoniae CWL029
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 10421)
AUTHORS Kaiman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R.W.,
Olinger, L., Greenwood, J., Davis, R.W. and Stephens, R.S.
TITLE Comparative genomes of Chlamydia pneumoniae and C. trachomatis
JOURNAL Nat. genet. 21 (4), 385-389 (1999)
MEDLINE 9920606
PUBMED 10192388
REFERENCE 2 (bases 1 to 10421)
```

AUTHORS Kalman,S., Mitchell,W., Marathe,R., Lammel,C., Fan,J., Olinger,L.,
Grimwood,J., Davis,R.W. and Stephens,R.S.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1998) Program in Infectious Diseases, University
of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
FEATURES Location/Qualifiers
source 1..10421
/organism="Chlamydochila pneumoniae CWL029"
/mol_type="genomic DNA"
/strain="CWL029"
/db_xref="taxon:115713"
133..2751
/gene="CPn0727"
133..2751
/gene="CPn0727"
/codon_start=1
/transl_table=11
/product="CT619 hypothetical protein"
/protein_id="AAD18866.1"
/db_xref="GI:4377032"
/translation="MSTFSTQNRRTISGSTRILKDHKYSQGFDPSPVAINLEELN
SGIYALHMLNAGSEINVAALNPNTTPTTSTWTDYKHSRPPQSSPRAPSQTPT
DIVSAALALVLDVGLAELVASVTEIDLGALSTISTVQLMASYGLTLTAQOK
WFSVSVPEKLEHVKQKAAEIOAKOBEIKAVLEAKGVSTEEIALIKLEVPDVI
AADPFKEFIEPLHTYRAKYGAPIQEMNENAIOLPTPEALTPDNVNEVMGTLSI
LOALDDAIKAPALGGQOEIITILQTLVPLVDKTFKAEFDLIYATQIPNTASLKL
YLDRQIAETRGKITVQNSIQNLSRQKVVNNRSMLETQLSMFOQAQNCVFVWIS
QANALNATNKYISAVLTSMENYGLLCLSYMYERLADDERKAFQSNVEXPLPHI
VGGSWNGWIKMAAYCELAEYSLGTAVSQDQIKAYLQTRGNEFKATRHFFHNGD
QMYQFANETVGNCLTANGAIQDPLGFTREAMTVGTVEADVNAQAQILNEFNFA
ATAHLQLOLQIAEQLOKADLDLPGKASTENKFAVAMITSESLGDLISMLNSQ
LPKQEAFLKLEIBENFNLAALNLSLIQITNEFTSTVYSLSSVVSQSKQNL
AGDYELLAAAREEYIYDRTAKQALINVLNGLQKLSLPGATSAQKQEMLNAT
YYQSVLNTNLGLVLESLAGLQMTLQTSNNKYDKSVFKIESFDWIPTLALESF
LTSQFNISATGLPLFTQVSDQQTYSQGTQQLNLQNMQTTIQEWLTVSTMSQ
VLNGLISQAGIYSN"
complement (3078..5033)
/gene="CPn0728"
complement (3078..5033)
/gene="CPn0728"
/codon_start=1
/transl_table=11
/product="ChlPn 76 kDa Homolog 1 (CT622)"
/protein_id="AAD18867.1"
/db_xref="GI:4377033"
/translation="MWNPIGGPIDEITERTPPDLQAQLEAAGNKAASABAQRIAGAE
AKPKESKTSVERMSILRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPT
PPTFDYDQTAQYATYITFTSLADIQAALVSLQDAVNIKDITATDEBTAIAEWM
ETKNADAKVGAQITELAKVASDNQALDLSGLKLTSPDLLQAALQSVANNKAAELL
KEMQDNPVPGKTPAIAGSIVDQDQATQIEKQGNARIDAYFAGONASGAVENAKSN
NSISNIDSAKATATQIAEQAKKFPDPIQEAQOMVIOAEKDLKNIKPADGSDV
PNFTTVGGSKQSSIGSRVSMLLDAENETASILMSFRQMIHNFENTENPDQAA
QBELAQAARAAAGDSDAAALADAQALEAALGKAGQQQGLNALGQIAASAVWSA
GVFPAAASSTGSSVKQIKTSKTSQSDYKTQISAGYDAYKINDAYGRANDATRVI
NNVSTPALTRVPARTEARPEKTDQALARVTSNRTGLGVYQSVSALQSVQVMIQ
SNPQANNEIRQKLTSAVTPKPPQPGYFVQLSNDSTQKFIKULESLFAEGSRFAAEIK
ALSFTNSLFIQVLNIGLSYGLQ"
complement (5145..6413)
/gene="CPn0729"
complement (5145..6413)
/gene="CPn0729"
/codon_start=1
/transl_table=11
/product="ChlPn 76 kDa Homolog 2 (CT623)"
/protein_id="AAD18868.1"
/db_xref="GI:4377034"
/translation="MKQVYQWLASVVLALTISGYAELPLSEOKVKSHYTTILDEVK
DYLKRGVETRKQDGLVIRAGDVRAEWLVEREDIKNPSDKQKYNPLVNNYSEFVL
YIDYABRNWLSKKNWTATAGENTAAGVDINRAFLGYRFYKNPETRTDFPMEIGRS
GLDLPSEVQVQFGLGHIYTRLSKDYPIQYIVHGGFFVNNMTKHYAVWVEGI
LNRLPKQFVFKCSVVDNFTVPSTSTTEKAATNAMKYCYWQWLMKHSQVFWING
QKPELYLYGAPLAWKTLNGLKENLAWFTGGTGLGLRKGADWSATVRYEYVA
LSVPEIDVSGIRGNLLKFWFAQIAAENYDPEKANGFTNYKGFSAIYWGITDLSLSPR

gene
complement (6516..8159)
/gene="mvn1"
/notes="synonym: CPn0730"
complement (6516..8159)
/gene="mvn1"
/codon_start=1
/transl_table=11
/product="Integral Membrane Protein"
/protein_id="AAD18869.1"
/db_xref="GI:4377035"
/translation="MSRKNEVSLABSNIENILSGTSCSRITGIFERRIAMATYFADPI
VAAPWLGRTVFFLRKILGGLILEQAFIHFELRAQSLDRAAFFRFRRLKGSTI
IFTLLIEAVLWVVLQVEGTDMILLTLLPCGIFLMYNNVNGALLHCENKFPVGV
LAPVYVNIWIFFVIAARSDPRRIIGLSVALVIGFFEWLITVGVWKFLEAKSP
PQHSVRALLAPLISITSSIQNLNLSIDICLARYVHEIGPLYIMSLIKYQLPFIH
LFGGVFTVLLPAISRCVQDERHGLKMKFVLTITMSVMIIMTAGLLIALPVGVRV
LYEHGFLPOSANVAIVRVLRGCGASIIIPMALPLVSLVFAQRCYVPLFIFIGTALA
NIVLSVLGRWLKQVSGISYATISITAWQLYFWMYSSKRLPMYSKLWESRRSIK
VMGTTMLACMITLGLNLIQTYYVIFINLPTPLAWPLSSITAQAIAPLSESCIFLAPL
YGFAKLLRVEDLILNLASFEYWRGQGLLQQRHYMQDTQN"
8143..8319
/gene="CPn0731"
8143..8319
/gene="CPn0731"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAD18870.1"
/db_xref="GI:4377036"
/translation="MSFLIKPSPPLKQDYLFDISPYTSSITIGGSYFKLNKASL
QSTLRLRSISILS"
8654..9535
/gene="nfo"
/notes="synonym: CPn0732"
8654..9535
/gene="nfo"
/codon_start=1
/transl_table=11
/product="Endonuclease IV"
/protein_id="AAD18871.1"
/db_xref="GI:4377037"
/translation="MKVLPPSPILLGAHTSTAGLKNAIYEGRDIGASTVQLFTANQ
RQWRALKEVIEDFAKLETDLSYIMSHAGLYINGPADPVPVILEKSRIGIYQEL
DCITLGSFVNFHPHGAALKSKEDKMKIVSPSQSAPLDPSPPLVLLTETTAQBT
LIGNSFEELGYLQNLKNOIRICVCDVTCHIFAAGYDITSPQGWEDVLENFDFYGLS
YLAFNLNDSMPLGANKORHAPLGRGYIGKESFKELMTDTRKIKYLETPGGEN
WQKEIGELLKFSKNRDS"
complement (9654..10283)
/gene="rsa"
/notes="synonym: CPn0733"
complement (9654..10283)
/gene="rsa"
/codon_start=1
/transl_table=11
/product="S4 Ribosomal Protein"
/protein_id="AAD18872.1"
/db_xref="GI:4377038"
/translation="MARYQCKNKNRVARFRGANIFGRSRNPLKKKPHPPGQGMQRKKK
SDYGLQLEBKQKLCACYMKEQVLKAFKEVTHKQNVQAFMEFLERFECRLDNVYRM
GPAKTIFAAQQLVAHGHIILVNGRVDRRSFFELRPGMQISLKEKSKLQSVKDALESKD
ESSLPYSISLDDTKGFKGELLVSPQDIEAQLPLPLINISVWCEFLSHRI"

Query Match 99.9%; Score 1455; DB 1; Length 10421;
Best Local Similarity 100.0%; Pred. No. 3.2e-301;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATAAATCTTTAAAAACAGCTCGCATTAATTAATTATAGTAGAGACTTTTATTTTATTTT 60
Db 5133 ATAAATCTTTAAAAACAGCTCGCATTAATTAATTATAGTAGAGACTTTTATTTTATTTT 5074
Qy 61 ATAATAAACTAAAAGATTTTATTTATTTTTCAGITTTTATGTTTAACTCTATTGCTCC 120


```

DRSHYTPERANLIWHPDAFDNDLLISDIKRLKNNNEIQAQVDFVLGNRSKTEJET
IYPSKVLIVFENQELRDLMDIRIFVDTDADERILRMVRDVBQEQDSVDCIMS
RVLSMVKPMHEKIEPTKVDYIIVHGVRQNVVTNLSQIKNHNLENALESDETVM
VNSK"
/gene="CP0012" .2021)
complement(1047. .2021)
CDS
/gene="CP0012" .2021)
complement(1047. .2021)
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37908.1"
/db_xref="GI:7188951"
/transl_table="MERKYYALAYYYITRDNPHEEIALHKLFLELDVSCRIVYEQ
GINDQSGYSPHAEVLMOWLKERPNFSKIKFKIHKENTFPRIITVKYKELAAIGCE
VDLSQAKHTISPOEMHEKLOENRCLLLDVNNVEMKIGHEFNATLPDIOTFRFFPEYA
EKLAQECDDPTTVMYCTGIGRCELSPVLEKGFKEVQLDGGVIAIYQQVGTGKW
LQGLVFEDRLAIPIDESDPVPIAECCHCQTPSDAYNCANTDNCALFLCCDCIH
QHQCCGECGQSQFVRKFDPSRGKPFRAHLCEISENSESASCCLI"
2238. .2867
/gene="CP0013"
2238. .2867
CDS
/gene="CP0013"
2238. .2867
/notes="similar to SP:P81289; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S4"
/protein_id="AAF37909.1"
/db_xref="GI:7188952"
/transl_table="MARYCYCPKNRVARRFQANIFGRSRNPLLKPPHPPGQHGQRKK
SDYGLQLEERQKQKACQYKIMEXQLVKAFKEVIHKQNVQAFMELEPERCELDNNVYM
GFAKTIQAAQLVAHGHIILVNGRVRDRRSPFLRPGMOISLKEKSKRLQSVKDALESD
ESSLPYSYISDKTGFKCELLVSPEDQIEAQLPLPINISVVCFLSHRT"
complement(2986. .3867)
/gene="CP0014"
complement(2986. .3867)
CDS
/gene="CP0014"
2238. .2867
/notes="similar to SP:P12638 GB:M22591 PID:146954
PID:405898 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="endonuclease IV"
/protein_id="AAF37910.1"
/db_xref="GI:7188953"
/transl_table="MKVLPPLPSIPLGAH"TAGGLKNAIYEGEDICASTVOIFRANQ
RQWRRAKBEVIEDPKAALKETDLSYIMSHAGYLINPGAPDPVILEKSRIGIYQEL
DCITLGLSFVNFHPGALKSKEDCMNKIVSSFSQAPLFDSPPLVLLLETTAGQT
LIGNFEEGLYVNLKNQIPIGVVDTCIHPAAGYDITSPQWEDVINEFDEYGLS
YLRAFHLNDSMFLGANKRHAFLGEGYIGKESFKMLTDERTRKIPKYLETEPGPEN
WQXIEGLLLKFSKNRDS"
complement(4202. .4378)
/gene="CP0015"
complement(4202. .4378)
CDS
/gene="CP0015"
2238. .2867
/notes="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF37911.1"
/db_xref="GI:7188954"
/transl_table="MSPLLKRPSPPLKQDYLFDISPYTSSBITIGSGSYFLKNKASL
QSSTLRISISIIIS"
4362. .6005
/gene="CP0016"
4362. .6005
CDS
/gene="CP0016"
2238. .2867
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"

```

```

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37912.1"
/db_xref="GI:7188955"
/transl_table="MSKKNESVLSARSINILSGTSCSRITGIFREIMATYEGADPI
VAAFNLGFRVFFRLKILGLLEQAFIPHEFLRAQSDLRDAFFFRFRLKGGSTI
ITFLIEAVLWVVOYEEGTDMILLTLLPCGIFLMMYNNVAGALLCENKFPVGV
LAPVVNIIWIFVIAARSDPRERIGLSVALVIGFFEWLITVPGVMKFLKAKSP
POBDSVRALLAPLSGILTSIFOLNLLSDICLARYVHEIGLYPLYSKLYQLPPIH
LFGSGVFTLLPAISRCVOREOHERGLKMKVLTLTMSMIITWAGLLILLAPGVRV
LYEHGLFPQSAVYAVVRVLGCGASIIIPALAPLVSVLPFAQOYAVPIFGCTALA
NIVLSVLGRWLKDVSGISYATISAWVOLYPLWYSSKRLPMYSSKLLWESRRSITK
VMGTMLACMITGLNLTITTVIFNLPLTFLAWPLSLITQAIAFLSESCIFLAF
FQFALLRVEDLINLASFYWRGQRLQORQHVMDTON"
6108. .7376
/gene="CP0017"
6108. .7376
CDS
/gene="CP0017"
2238. .2867
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37913.1"
/db_xref="GI:7188956"
/transl_table="MKQVYQMLASVVLALITISYAEPLSEQKYSHTYTLDEVK
DYLKRGFVETRKQDGLVRIAGVRARWLYFREDIKNPSDKKYNPLPNRYRSEFVL
YIDYRAERNWLSKNNMTAIGAGTAAGVDINRAFLGYRKYKPNRTRTDFNWEIGRS
GLGDLPESEVQSQSNFQDGLHYWTRPRLSKDYQVILVHGSPVNNKHYAMVBSGI
LNRLPKQFVKSVVDWNTVFVSESTTEKATNANKYKVCWQNLVGHKHSQVPWING
QKRPPLYGAFLMNLAKATKTLNKENLAWPIGGTLGGLRKGAGWSAFVRYEYEA
LVSPEIDVSGIGRGNLKFQFAQIAANYDPKEANGFTNYKGSALYMYGITDLSLFR
AYGAYSKPANDKLSGDTFRKFDLGIISAF"
7488. .9443
/gene="CP0018"
7488. .9443
CDS
/gene="CP0018"
2238. .2867
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37914.1"
/db_xref="GI:7188957"
/transl_table="MVNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQIAGAE
AKPKESKTSVERWSILRGAVNALMSLADKLGIASSNSSSTSRSDVDTATATPTP
PPTFDYKTOQACTAYDTIFTSTSLADIOALVSLQDANTNIDKTAATDEETAIAEAW
ETKNADAVKYGQITELAKYASDNQAILSLGKLTSLDILLOALLOSANNKAELL
KMQDPVYVFGKTPALAQSLVDQTDATATQIEKGNARIDAFAGQASGAVENAKSN
NLSNIDSAAKAIATARTQIAEAKKFPDPSILOEAMQVIQAEKDLKNIKPADSDV
PNFTVTGSGKQSGSIGIRYVSMLLDAENETASTIGSFQMTIMFTNPNDSQAA
QBELAAQAAKAAAGDSSAAALADAKALAKGAKQOQGLNALGOTASAAYVSA
GYVPAASISGSSVKOLYKTSKTSQDYKTOISAGDAYKSINDAYGRANDATROVI
NNVSTPATESVPRATEARGPEKTDQALRVISGNSRTLGDVYSOVSALQSMQILIQ
SNPQANNEIRKLTSAVTKPQFGFYVYVNDSTQKFIKLESFASGSRPAEIK
AUSFETNSLFIQQVLVNLISGLSYGLQ"
complement(9770. .12388)
/gene="CP0019"
complement(9770. .12388)
CDS
/gene="CP0019"
2238. .2867
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37915.1"
/db_xref="GI:7188958"
/transl_table="MSTFSIQNRLRTISGESTRIIKLDHKYSGFDPSPVPAINLEELN
SITVALRHLMLNAGQENTNVAALNPNITFTTSTWTDYKHSRPOASSPRASSQPTT
DIVSAAALALVIDGGLAEVASTVEIDGLASTISTIRQLMASYLGITLTAEQEK
VWFSYSPSEKNNLHVKQEAARTQAKERIKAVLEAKGVSTEEIALLKEYPDYI
ADDFKEFIEEPLHTYRAKVAPIQEMENALQLLTPPALTIDNVNNGNTLSTI

```

Query Match		99.9%; Score 1455; DB 1; Length 12548;
Best Local Similarity		100.0%; Pred. No. 3.1e-301;
Matches 1455; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	1	ATAAATCTTTTAAACACAGCTCGCATTAATTAATAGTGAGAGCTTTTTTTTATTTT 60
DB	7388	ATAAATCTTTTAAACACAGCTCGCATTAATTAATAGTGAGAGCTTTTTTTTATTTT 7447
QY	61	ATAAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
DB	7448	ATAAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 7507
QY	121	AGTCTCTATAGACGAACAGACGACCTCCGAGATCTTTCTGCTCAAGGATTTGA 180
DB	7508	AGTCTCTATAGACGAACAGACGACCTCCGAGATCTTTCTGCTCAAGGATTTGA 7567
QY	181	GGCGAGTCAGCAATAAGAGTCCGGAAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC 240
DB	7568	GGCGAGTCAGCAATAAGAGTCCGGAAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC 7627
QY	241	TAAAGAACTTAAGACCGATTTCTAGAGGATGAGGACATCTTCCGTTCTGCAAGTGAATGC 300
DB	7628	TAAAGAACTTAAGACCGATTTCTAGAGGATGAGGACATCTTCCGTTCTGCAAGTGAATGC 7687
QY	301	TCTCATGAGTCGGCAGATAAGCTGGGTATTTCTTAGTAACAGCTCTCTTCTACTAG 360
DB	7688	TCTCATGAGTCGGCAGATAAGCTGGGTATTTCTTAGTAACAGCTCTCTTCTACTAG 7747
QY	361	CAGATCTGAGAGCTGAGCTCAACGACGACGACGACCTAGCCCTCCACCCACGTT 420
DB	7748	CAGATCTGAGAGCTGAGCTCAACGACGACGACGACCTAGCCCTCCACCCACGTT 7807
QY	421	TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACT 480
DB	7808	TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACT 7867
QY	481	AGCTGACATACAGGCTGCTTTGTGAGCTCCAGGATGCTGCTCACTAATAAAGGATAC 540
DB	7868	AGCTGACATACAGGCTGCTTTGTGAGCTCCAGGATGCTGCTCACTAATAAAGGATAC 7927
QY	541	AGCGGCTACTGATGAGGAACCGCAATCGCTCGGAGTGGAACTAAGAAATCCCGATGC 600
DB	7928	AGCGGCTACTGATGAGGAACCGCAATCGCTCGGAGTGGAACTAAGAAATCCCGATGC 7987
QY	601	AGTTAAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTTCGGATAACCAACGAT 660
DB	7988	AGTTAAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTTCGGATAACCAACGAT 8047
QY	661	TCTTGACTCTTTAGGTAAAGTACTCTTCCGACCTTTACAGGCTGCTTCTTCCAAATC 720
DB	8048	TCTTGACTCTTTAGGTAAAGTACTCTTCCGACCTTTACAGGCTGCTTCTTCCAAATC 8107
QY	721	TGTAGCAAAACAATAACAGCAGCTGAGCTTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
DB	8108	TGTAGCAAAACAATAACAGCAGCTGAGCTTTCTTAAAGAGATGCAAGATAACCCAGTAGT 8167
QY	781	CCAGGGAACCGCTGCAATTTGCTCAATTTTAGTTGATCAGACAGATGCTACAGCAC 840
DB	8168	CCAGGGAACCGCTGCAATTTGCTCAATTTTAGTTGATCAGACAGATGCTACAGCAC 8227
QY	841	ACAGATAGAAAGATGGAATGCGATAGGATGCAATTTTTCAGGACAGACGCTAG 900
DB	8228	ACAGATAGAAAGATGGAATGCGATAGGATGCAATTTTTCAGGACAGACGCTAG 8287
QY	901	TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACGTTAAGC 960

DB	8288	TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTGAGCTAAAGC 8347
QY	961	AGCAATCCCTACTGCTAGACACAAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 1020
DB	8348	AGCAATCCCTACTGCTAGACACAAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 8407
QY	1021	AATCTTCAAGAGCGGAAACAAATGGTATATACAGCTGAGAAAGATCTTAAATAATATCAA 1080
DB	8408	AATCTTCAAGAGCGGAAACAAATGGTATATACAGCTGAGAAAGATCTTAAATAATATCAA 8467
QY	1081	ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAACAACA 1140
DB	8468	ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAACAACA 8527
QY	1141	AGGAAGTAGTAGTGTAGTATTTCGTGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGAC 1200
DB	8528	AGGAAGTAGTAGTGTAGTATTTCGTGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGAC 8587
QY	1201	CGCTTCCATTTTGTGTTCTGGGTTTCGTGATGATGATGATGATGATGATGATGATGATGATG 1260
DB	8588	CGCTTCCATTTTGTGTTCTGGGTTTCGTGATGATGATGATGATGATGATGATGATGATGATG 8647
QY	1261	TGATTTCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAAGCTAGAGCAGCAAAAGCGCTGG 1320
DB	8648	TGATTTCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAAGCTAGAGCAGCAAAAGCGCTGG 8707
QY	1321	AGATGACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 1380
DB	8708	AGATGACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGG 8767
QY	1381	TAAAGCTGGCAACCAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
DB	8768	TAAAGCTGGCAACCAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 8827
QY	1441	TGTGAGCGCAGGAGT 1455
DB	8828	TGTGAGCGCAGGAGT 8842
RESULT 6		
AE017159/c		
LOCUS		300066 bp DNA linear BCT 25-JUL-2003
DEFINITION		Chlamydomphila pneumoniae TW-183, section 3 of 4 of the complete genome.
ACCESSION		AE017159 AE009440
VERSION		AE017159.1 GI:33236383
KEYWORDS		
SOURCE		Chlamydomphila pneumoniae TW-183
ORGANISM		Chlamydomphila pneumoniae TW-183
REFERENCE		Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
AUTHORS		1 (bases 1 to 300066) Geng, M.M., Schumacher, A., Muehlendorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T., Essig, A., Marre, R. and Melchers, K.
TITLE		The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis
JOURNAL		Unpublished
REFERENCE		2 (bases 1 to 300066) Geng, M.M., Schumacher, A., Muehlendorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T. and Melchers, K.
AUTHORS		Direct Submission
TITLE		Submitted (28-MAY-2002) RDR/JT, RDR/FG, RDR/BT, ALTANA Pharma, Byk-Gulden-Str. 2, Konstanz 78467, Germany
JOURNAL		
FEATURES		Location/Qualifiers
source		1..300066 /organism="Chlamydomphila pneumoniae TW-183" /mol_type="Genomic DNA" /strain="TW-183" /db_xref="taxon:182082" /note="type strain biovar: TWAR" 126..797
gene		

```

/locus_tag="CpB0543"
126..797
/notes_tag="CpB0543"
/locus_tag="CpB0543"
/notes="similar to hypothetical protein CT433 - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98472.1"
/db_xref="GI:33236384"
/translators="MILYGLNCKTARKYQAHYLPILTPPYAKSPONKRALQFLPQ
ATHVILTPSGSTHLFSMTSLSKATLTKTKVLCIGESTKERLISFLGQVKVAVATQ
EIAEGLPDLQALPSSARILYPHSSARVIRELYNRPFFSPHYTYKPKLXNNI
LSYKXKILLSPSTVRPAKILFRPFPEKTYWCQORMLQEFQKFSQKQVSLLETGK
SRVSP"
862..1194
/locus_tag="CpB0544"
862..1194
/locus_tag="CpB0544"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98473.1"
/db_xref="GI:33236385"
/translators="MSSATPFGDGTAPSLFPLATPRYNFKLALFTVIALVWIAL
IATIALGLCHPLCSFIFTALPLFISYICSHVARNVYIALDVVDHSLQDMRS
HSPIFSDR"
1242..2321
/locus_tag="CpB0545"
1242..2321
/locus_tag="CpB0545"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98474.1"
/db_xref="GI:33236386"
/translators="MSPGSTESSQVSLSVYPRDKETAPKKOFTIAKISTLAILASL
ALGAVAGISLTIVLGNVPLALLITLPSVTVFVYHMTSKVSNQKVLQBNFK
PLKANQKNCVDSYEMQYNHNLAPKFKVALQDASQFFQFTLIGLURVIEKNQST
GIIFNVPGLINDINTATNLSTILYSLDKDSYMDCKQREGDPAKGEOPFSPTEVNR
VKLPNALDQTNLNLSSAEKKSILFTFLGHVCGPKSEELPNQOEYRQALLAYENCL
KAAIESHAALPLFTSVYVEPPEILPKEGTFYWDNQQAFCRALLDAIQNTAIR
YQPSLLVILQDPFNTIESQSRSEE"
2661..3425
/locus_tag="CpB0546"
2661..3425
/locus_tag="CpB0546"
/notes="similar to hypothetical protein CT398 - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98475.1"
/db_xref="GI:33236387"
/translators="MHDALLSILAIQELDIDKIMLRVKKHOKELAKVQSLKSDIRR
KVQEKLEMLNATQIRDGNRIQEISEQINKLENQAAVKKMDENALTOENTANK
ERESLHQSLDKQAGEDDLIVSLKESLASTENSSSVIEKEIFESIKKINEGKAL
LEORTLKHATNPILLSYTERLNNKDRVVFEIKRVCSGGCHIVLTPOHENLVRRKD
RLIFBHCGRILYQWESQVNAQENSAKRRRRRAV"
complement(3846..4835)
/locus_tag="CpB0547"
/locus_tag="CpB0547"
/notes="kpsF"
complement(3846..4835)
/locus_tag="CpB0547"
/locus_tag="CpB0547"
/notes="similar to kpsF protein - Escherichia coli"
/codon_start=1
/transl_table=11
/product="kpsF"
/protein_id="AAP98476.1"
/db_xref="GI:33236388"
/translators="MPSPMISTDVCODILGKQKEAVDFFQAPQPKEMQLAEKILGH
SGWVFTSGVKGSCVARKVATLQSLSERALFFSPVDLLHLDGLVSPGDIVCLFSKS

```

```

GETQELLDTVPHLKSRRAILVAITSMPYNSNLAALSDLVILPSVABELDFNLPITNST
TCQMTFGDFLAMLHFSRGSVLSITYGKHPSSQVGMKANGKVDKDFPFPKTEFVFCHLG
DKVFSVEFSAYGCCVCIVDPQFRLMGIFTDGLDRSLASGGSVLSISLEKVMTA
NPRCITFEDSDIAIALQOLMESSSPVAVLPVLDDNEENRHVTGLLHMHTLAKAGLL"
complement(4862..6043)
/locus_tag="CpB0548"
complement(4862..6043)
/locus_tag="CpB0548"
/notes="similar to dihydrolipoamide S-acetyltransferase - Bacillus stearothermophilus"
/codon_start=1
/transl_table=11
/product="dihydrolipoamide S-acetyltransferase"
/protein_id="AAP98477.1"
/db_xref="GI:33236389"
/translators="MIFEPFKIGTSSGSSIVRWLKNLGDHVARDEPLIEVSTDKI
ATELPSPKAGRLVRFVNEGDEVAAGDVLGLIELEISEADDESTCPLTSCSTKSEA
GSSSSVWSPFSPAVLSLAQREGIDLNLOKIAGTGGKGRVTRQDLRAVISESQVSIPE
IFQGEVNRIPMSPPLRAIASLSKSSDEVPHASLVVDVTDLMHLISERQRFLOTH
GVKLLTSFVQCLAQTLQFPELLSGDGTIIVMKKSVNVGAVNLNKEGVVPPVTH
NQDRLGVSIAKALADLSRRLNKLDSEVDQGSVTNFGMTGALIGMPIIRYPEV
AILGICTIQKRVVVRDDDSLAVRKMVYVTLTFDHRVLDGIYSGSEFLTSLKNLESVTM
G"
complement(6057..7301)
/locus_tag="CpB0549"
complement(6057..7301)
/locus_tag="CpB0549"
/notes="similar to proton/sodium glutamate symport transporter family"
/codon_start=1
/transl_table=11
/product="proton/sodium glutamate symport transporter"
/protein_id="AAP98478.1"
/db_xref="GI:33236390"
/translators="MKLMKIFIGLVFGVTGLVLEDKAIFFKPIGDIENLLSMVVY
PIVFCSMVLGIASISDMKKLGISVGLVLTALAVIGLCPAWIFSPGNGCQDFA
QAQMSDVAVIDSNKTAAYFSLIAQVPSNPVRSFAGNLIQIIFALFIQIALRL
SERGPVRFIDGSEIMLRMNWIMSPAPVGVGASMAWISGNHGLGVWLQKGLII
AYLACLFHATLVFGLVRFCKMSFSEKLSMMMDAISCAVSTASSATILPVTMRCVS
KILGVSAEVSFGVLPILGATVNMNGTAIFQGMAAVFAIQAYNCPCLSSLLLVVATF
SVAGGAGVGGGMITIGLSVASVGLPIQGIALLAGIDRLRDIVGTPMNIGDAVVATY
VASGELSPYESIKQESVETT"
complement(7301..8398)
/locus_tag="CpB0550"
/locus_tag="CpB0550"
complement(7301..8398)
/locus_tag="CpB0550"
/notes="similar to ycaH protein - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="YcaH"
/protein_id="AAP98479.1"
/db_xref="GI:33236391"
/translators="MKKRPSLTFLFYRRVTIAISLEGLMGWGLSLLSKVPAFLVA
CWRNFSWSPVBARSTVIVSGNIVVGGAKTPTVLALRALRGYSCVLSRGVSKQ
SRQKXILVYDSKVBSASVGDPELMAKLEPESVWVHKORRISAAARAEKGILL
DVLQVRSNAPQIFVFKPTIASVVTNGERIPKEALRELVRGVFCGLFPQGFGLMLRE
EGIHILGKYLPHDAITKKELNYFCQOMAMRQGGQLLCTEKDVKLPRLSGEVSLIP
IAKVMRLSVNODDTLSLNMLEQIHKNRN"
complement(8596..9393)
/locus_tag="CpB0551"
/locus_tag="CpB0551"
complement(8596..9393)
/locus_tag="CpB0551"
/notes="similar to tRNA methylase spoU - Escherichia coli"
/codon_start=1
/transl_table=11
/product="tRNA methylase"
/protein_id="AAP98480.1"
/db_xref="GI:33236392"

```


the United States
J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)
2029986
10839753
2
Shirai, M., Hirakawa, H., Kimoto, M., Takachi, M., Kishi, F., Ouchi, K.,
Shiba, T., Ishii, K., Hattori, M., Kuhara, S., and Nakazawa, T.
Comparison of whole genome sequences of *Chlamydia pneumoniae* J138
from Japan and CWL029 from USA
Nucleic Acids Res. 28 (12), 2311-2314 (2000)
20330349
10871362
3 (bases 1 to 300550)
Shirai, M.
Direct Submission
Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
School of Medicine, Department of Microbiology, 1-1-1
Minamikogushi, Ube, Yamaguchi 755-8505, Japan
(E-mail: mshirai@po.cc.yamaguchi-u.ac.jp, Tel: 81-836-22-2227,
Fax: 81-836-22-2415)
On or before Sep 15, 2000 this sequence version replaced
gi:6172398, gi:6172300, gi:6172396, gi:6172398, gi:8978889.
Location/Qualifiers
1. .300550
/organism="Chlamydia pneumoniae J138"
/mol_type="genomic DNA"
/strain="J138"
/db_xref="taxon:138677"
/note="synonym: Chlamydia pneumoniae (strain J138)"
complement (21. .794)
/gene="dap"
complement (21. .794)
/gene="dap"
/codon_start=1
/transl_table=11
/product="diaminopimelate epimerase"
/protein_id="BAA98725.1"
/db_xref="GI:8978890"
/translations="MAFYSPSTISKYIYSGAGNRLGTLPEVEDRFLCOETRVD
GFILXPSSCADAQLIIFNSDGRPMCGNLGKALAHASQKSDISVSTSDGLYS
GIFYMDRVLDMLADWRASVHRLSRPDLPEVVCVCIHTGVHAWILFEISTDLV
SILGFLRHQTEISPGVNVNFIQLGHCOLRVRTYRVEGGETAACGTGALASLV
SNSYGKESQIHTWEGELMTVSQNRGVYVQSGVTRDL"
complement (763. .1338)
/gene="clpP 1"
complement (763. .1338)
/gene="clpP 1"
/codon_start=1
/transl_table=11
/product="CLP protease"
/protein_id="BAA98726.1"
/db_xref="GI:8978891"
/translations="MADGEVHKURDIIEKLLRARRVFPSEPTKSAIDAIKKLWYL
ELDKPKPIVFINSGSGSDVAGFVMDQIKMLTSPVTVTGAAASGVSLSLCAAP
GRFAPTHLHPSQIGSGIPGQATDLDIHAREILKTKARIIDVVEATNQPDII
EKAIIDRMWMTANEADFGLLDILFSFNDL"
complement (1358. .2917)
/gene="glyA"
complement (1358. .2917)
/gene="glyA"
/codon_start=1
/transl_table=11
/product="serine hydroxymethyltransferase"
/protein_id="BAA98727.1"
/db_xref="GI:8978892"
/translations="MLKVFKEFKFAIVEIFTKWAVVLSLLHKLLENASGKQSLAS
TAYLALDHLNAPPSIGERIIDELKSQRSLHMIASNYSSLSQANMLITDKYC
EGSPFKRFYCCCNVDIAEWECETAKELPAACVOPHSGADANLAWLITHKY
GCPAVSKLGYKTVNLTETEEYTLKAMSSCVCLGSLASGGHLTGNVFLNWSKLM
RCPDYVNPTECFDVAIEIRLAKYKPKVLIAGYSISRRNPAVLKQIAEDCGSLV
WVDMHAPGAVAGVDFDEENPIFYADIVTITHTKLGRGQVNNARLAEFLSHGLRLTGT
CLPMGGPFLHVAATKVALKEALSVDFFKYAHQVNNARLAEFLSHGLRLTGT
DNHMMVITDLSGISGKIADILSSVGIANNRSLPSDATGKWDTSIGRLGTALTL

GMGIDMEERVADIIVKVLNIRLSCHVEGSSKNKGELPEALQAQEARDRVNRLLRFP
LYPEIDLEALV"
3107. .3778
/gene="CPj0522"
3107. .3778
/gene="CPj0522"
/codon_start=1
/transl_table=11
/product="Ct433 hypothetical protein"
/protein_id="BAA98728.1"
/db_xref="GI:8978893"
/translation="MTVLGLNQKARKYQAHVLPILTLFPYAKSPONKRALQFLPQ
ATHVILTPSSSTHLFLSLMGLSKATLKTATLCTIGESTKERLSFLGQVKKVATQ
EIAEGIFLLOALPSSARILYEHSSLARVIREFLYNRTFFSYPHYTVKPKLKNIL
LSKIKLIFTSPSTVRAFAKIPFPFKTYICQGRNLTQEFQKFSQKQVSLLETGLK
SRTSP"
3843. .4175
/gene="CPj0523"
3843. .4175
/gene="CPj0523"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98729.1"
/db_xref="GI:8978894"
/translation="MASSATPGDGTAPSLFPATEPRYNEKLALFVTLIALVIAL
IATTAIGLCIHPLCSFIFLTAIPFYISRYICSHYARNVYIALDVDPHSLKLDQRS
HSPIFSDF"
4223. .5302
/gene="CPj0524"
4223. .5302
/gene="CPj0524"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98730.1"
/db_xref="GI:8978895"
/translation="MSGPRTSSQVSVLYVPRDKEIAPKQKTIKTIALLIASL
ALGALVAGISLIVLGNPFLALLITLALFSVVTLVVHQTSSKVSNNQKVLQNFK
PUGKAEKNVCYINEMOFYNNHLNPKFKALQIDASQFQFTLTGLRVIEKQST
GIIFPGPTNIDNTATNLSTLYLTKDKSVWDTCQREGPAKGEPPFTVRV
VKLPALQDNLNLSAEEKSLPTFLHGVCGPKSELPNQOEYRQALLAYENCL
KAAISHAAIVALPLFTSVYVPEPILPKETGYMDNQTAFCRALLDAIQNTALR
YFQSRLLVILQDPFNTIESQSRSE"
5642. .6406
/gene="CPj0525"
5642. .6406
/gene="CPj0525"
/codon_start=1
/transl_table=11
/product="Ct398 hypothetical protein"
/protein_id="BAA98731.1"
/db_xref="GI:8978896"
/translation="MHDALLSLAIQELDKIMRLMRVKKHOKELAKVOSIKSDIRR
KYQCELEMLNKTQIRIDGENRIQIETSEINKLENQAAQKMDENALTOBMTANK
ERRSLEHSDLMDQAGGEDLIVLSKESLSTSSSVIEKEIPESIKKINEBGAL
LEQRTTELHATNPELLSIYERLLNKKORVVPIENRYCSCGCHIVLPQHENLVKRD
RLIFCHCSRILYWQESQVNAENSTAKRRRAAV"
complement (6827. .7816)
/gene="kpsF"
complement (6827. .7816)
/gene="kpsF"
/codon_start=1
/transl_table=11
/product="GutQ/kpsF family sugar-P isomerase"
/protein_id="BAA98732.1"
/db_xref="GI:8978897"
/translation="WPSMISTDVCQDILGKQKEAVDFPQAFQPEAKMLAEKILGH
SGWFPFGVKGSGCVARKLVAIQLSIRALFFSPVLLHGLGLVSPGDIIVLPSKS
GETQDLDTVPLKRRALLVAITSPYSNLAAALSDLVVILPVSVAELDFNLIPINST
TCQMTFGDPLAMLLFHSRGLSVLSTYKQHPGQVGMKANGKVPFKEFPFCHGL
DKVFSLEVSFVAGGCVIVDPQLMGIFTDGLRRSLASVGGSEVLSLEKYMTA
NPERCTEDSDIAIALQLMESSPVAVLVLDNEENRHTVGLLHMTLAKAGLL"

gene	complement (7843. .9024) /gene="sucB_2" complement (7843. .9024) /gene="sucB_2" /codon_start=1 /transl_table=11 /product="dihydrolipoamide succinyltransferase" /protein_id="BAA98733.1" /db_xref="GI:8978898" /translation="MIFEFPPKIGETSSGSIVRWLNKLDHVARDEPLIEVTDKI ATELPSKAGRLVFCVNEGDEVASGVILGILEEISEADDESTSCPLTSCPTKSEA GSSSSVWFSPAVLSAQREGIGLDNLQKTAGTGKGRVTPQDLEAVISBQOVSIBP IFQEVNRIPSPRRRAIASLSKSDSEVPHASIVVDVTDMLNLLSGRQRFLDTH GKUTITTSVFOCLAQRLRPFLNGLSDGTTIVMKSVNVGVAVNKNKSGVVPVLIH NCQDGLVISAALADISSRAINKLDPEVODGSVTVTFNFGMTGALIGMPIIRYPEV AILGIGITQKRVVVRDDSLAIRKMWYVTLTFDRVLDGIYSEFLLSLKNRLBSVTM G"	Db
CDS	217278 TGATGATTATAAGACTCAAGCGCAACAGCTTAGGATACTATTCTTACCTCAACATCACT 217219 481 AGCTGCATACAGCGTCTCTTTTGGTGAGCCCTCCAGGATGCTGTCACTAATAATAAGGATAC 540 217218 AGCTGCATACAGCGTCTCTTTTGGTGAGCCCTCCAGGATGCTGTCACTAATAATAAGGATAC 217159 541 AGCGGTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGATGCCGATGC 600 217158 AGCGGTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGATGCCGATGC 217099 601 AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 660 217098 AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 217039 661 TCTTGAATCTTTTAGGTAAATCTGATCTCTTCGACCTCTTACAGGCTCTCTTCTTCCAAATC 720 217038 TCTTGAATCTTTTAGGTAAATCTGATCTCTTCGACCTCTTACAGGCTCTCTTCTTCCAAATC 216979 721 TGTAGCAACAATAAACAAGCAGCTGAGCTTCTTAAAGAGATCAAGATCAACCCAGTAGT 780 216978 TGTAGCAACAATAAACAAGCAGCTGAGCTTCTTAAAGAGATCAAGATCAACCCAGTAGT 216919 781 CCCAGGAAAAACGCTCGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840 216918 CCCAGGAAAAACGCTCGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 216859 841 ACAGATAGAGAAAGATGGAATCGGATAGGATGCAATATTTTCAGGACAGAACCGTAG 900 216858 ACAGATAGAGAAAGATGGAATCGGATAGGATGCAATATTTTCAGGACAGAACCGTAG 216799 901 TGGAGCTGTAGAAAATCTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGC 960 216798 TGGAGCTGTAGAAAATCTAATCTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGC 216739 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 1020 216738 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 216679 1021 AATTCTTCAAGACGGGACCAAAATGGAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080 216678 AATTCTTCAAGACGGGACCAAAATGGAATACAGGCTGAGAAAGATCTTAAAAATATCAA 216619 1081 ACCTGCAGATGGTTCTCATGTTCCAAATCCAGGAACCTACAGTTCGAGGCTCCAAAGCAACA 1140 216618 ACCTGCAGATGGTTCTCATGTTCCAAATCCAGGAACCTACAGTTCGAGGCTCCAAAGCAACA 216559 1141 AGGAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 1200 216558 AGGAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 216499 1201 CGCTTCCATTTTGTATGCTGGTTTCGTGATGATTCAGATGATTCAGATGATTCAGATGATTC 1260 216498 CGCTTCCATTTTGTATGCTGGTTTCGTGATGATTCAGATGATTCAGATGATTCAGATGATTC 216439 1261 TGATTTCTCAAGCTCCCAACAGGAGCTCGCAGCACAAGCTAGACGCGAAAGCGCGTGG 1320 216438 TGATTTCTCAAGCTCCCAACAGGAGCTCGCAGCACAAGCTAGACGCGAAAGCGCGTGG 216379 1321 AGATGACAGTCTGCTGCTGAGCGCTGGCAGATGCTTCAGAAAGCTTTAGAACGCGCTCTAGG 1380 216378 AGATGACAGTCTGCTGCTGAGCGCTGGCAGATGCTTCAGAAAGCTTTAGAACGCGCTCTAGG 216319 1381 TAAAGCTGGGCAACAACAGGCGTACTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT 1440 216318 TAAAGCTGGGCAACAACAGGCGTACTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGT 216259 1441 TGTGAGCGCAGGAGT 1455 216258 TGTGAGCGCAGGAGT 216244	Qy
gene	complement (10282. .11379) /gene="ycaH" complement (10282. .11379) /gene="ycaH"	Db
CDS	99.9%; Score 1455; DB 1; Length 300550; Best Local Similarity 100.0%; Pred. No. 2.2e-301; Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 1 ATAAATCTTTAAACAGCTCGCAATTAATTAGTGAGAGCTTTTTTTTATTTTTT 60 217638 ATAAATCTTTAAACAGCTCGCAATTAATTAGTGAGAGCTTTTTTTTATTTTTT 217639 61 ATAAATAAACTAAAGATTTTATTTATTTTGGTTTATGTTAATCCTATTGGTCC 120 217638 ATAAATAAACTAAAGATTTTATTTATTTTGGTTTATGTTAATCCTATTGGTCC 217579 121 AGGTCTATAGCAAGAACAGACACACCTCCGAGATCTTTCTGCTCAAGGATTGGA 180 217578 AGGTCTATAGCAAGAACAGACACACCTCCGAGATCTTTCTGCTCAAGGATTGGA 217519 181 GCGGAGTCGAGCAATAAGAGTGGGAGCTCAAGAAATAGCAGTGGGAGGCTAAGCC 240 217518 GCGGAGTCGAGCAATAAGAGTGGGAGCTCAAGAAATAGCAGTGGGAGGCTAAGCC 217459 241 TAAAGAATCTAAGACCGATTCTGTAGCGGATGGAGCATCTTTCGTTCTGCAAGTGAATGC 300 217458 TAAAGAATCTAAGACCGATTCTGTAGCGGATGGAGCATCTTTCGTTCTGCAAGTGAATGC 217399 301 TCTCATGATCTGGCAGATAAAGCTGGGATTTCCTTAGTAACAGCTCGTCTTACTAG 360 217398 TCTCATGATCTGGCAGATAAAGCTGGGATTTCCTTAGTAACAGCTCGTCTTACTAG 217339 361 CAGATCTGACAGCTGGATCTCAACGACAGCAGCAGCAGCTAGCGCTCTCTCCACCCAGGTT 420 217338 CAGATCTGACAGCTGGATCTCAACGACAGCAGCAGCAGCTAGCGCTCTCTCCACCCAGGTT 217279 421 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 480	Qy

RESULT 8
CHT76KDA

LOCUS	CHT76KDA	2148 bp	DNA	linear	BCT 17-MAR-1994			
DEFINITION	Chlamydia pneumoniae 76 kDa protein gene, complete cds.							
ACCESSION	L23921							
VERSION	L23921.1	GI:435961						
KEYWORDS	76 kDa protein.							
SOURCE	Chlamydia pneumoniae							
ORGANISM	Chlamydia pneumoniae							
REFERENCE	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia. 1 (bases 1 to 2148).							
AUTHORS	Perez Melgosa, M., Kuo, C.C. and Campbell, L.A.							
TITLE	Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-kilodalton protein containing a species-specific epitope							
JOURNAL	Infect. Immun. 62 (3), 880-886 (1994)							
MEDLINE	94156481							
PUBMED	7509320							
COMMENT	source text: Chlamydia pneumoniae DNA.							
FEATURES	Location/Qualifiers							
source	1..2148							
	/organism="Chlamydia pneumoniae"							
	/mol_type="genomic DNA"							
	/db_xref="taxon:83558"							
CDS	1..2148							
	/codon_start=1							
	/transl_table=11							
	/product="76 kDa protein"							
	/protein_id="AAA23117.1"							
	/db_xref="GI:435962"							
	/translation="MTKKHVAWVVEGILNLPKQFFVKCSVDWNTFVPSSTTEKA ATNAKYKVCWMLQKHSQVPWINGOKPLYLIGAFVWLNPLAKTFTLNGKENIA WFLIGTGLGRKGDWATVREYVEALSVIDVSGIRGNLLKWFPAQJANVDP KEANFNKYRGSALYMYGTFIDLSFRAYGAYSKPANDKLSGDFTEFKDLGISAGF KSNFKLIFKRLALIISEFFFIYNTKTRPLLFEEFLVNPFGPGIDETERTPPAD LSAGLESAANKSAEQRJAGABAPKESKTDSEVERSTILRSVNALMSLADKLGA SNSSSSTRSADVSTATAPTPPPPTFDYKTQATYDITFTSLADIQAALVS LQDAVNTKTAADDETAAEWETKNADAVKVAQITELAKVSDNQALDLSGLK TSPDLLQAALLSVANNKAEILLKEMDNVPVPGKTPAIAQSLVDQTDATATQIEKD GNARDFAYQGRASGAVENAKSNSISNIDSAAKAAITAKTQIAEAKQFPSPILQ EAGMVUZAQKDLNPNPQSDVFNFTVGGSKQSGISIRVSLMDLDAENETFA SILMSGFRIHMFENTPDSQAQQLAAQRAAKAAGDDSDSAAALADAQKALEAL GKAGQQQGIILNALGQIASAAVSVGLPLQQL"							
ORIGIN	Query Match 99.2%; Score 1444; DB 1; Length 2148; Best Local Similarity 99.9%; Pred. No. 8.6e-299; Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;							
QY	1	ATAAAATCTTTAAAAACAGGCTCGCATTAAATTATTAGTGAGAGCTTTTATTTTATTTT	60					
DB	668	ATAAAATCTTTAAAAACAGGCTCGCATTAAATTATTAGTGAGAGCTTTTATTTTATTTT	727					
QY	61	ATAATAAACTAAAGAATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	119					
DB	728	ATAATAAACTAAAGAATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	787					
QY	120	CAGTCTCTATAGAGAAACAGACGACACTCCCGCAGATCTTCTGCTCAGGATTGG	179					
DB	788	CAGTCTCTATAGAGAAACAGACGACACTCCCGCAGATCTTCTGCTCAGGATTGG	847					
QY	180	AGCGAGTGCAGCAAAATAAGAGTGCAGGAGCTCAAGAATAGCAGGTGCGGAAGCTAAGC	239					
DB	848	AGCGAGTGCAGCAAAATAAGAGTGCAGGAGCTCAAGAATAGCAGGTGCGGAAGCTAAGC	907					
QY	240	CTAAGAATCTAAGACCGATTCTGTAGAGGATGGAGCACTTCGCTTGTGAGTGAATG	299					
DB	908	CTAAGAATCTAAGACCGATTCTGTAGAGGATGGAGCACTTCGCTTGTGAGTGAATG	967					
QY	300	CTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAACAGCTCGCTTCTACTA	359					
DB	968	CTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAACAGCTCGCTTCTACTA	1027					
QY	360	GCAGATCTGAGAGTGGATCTCAACGACAGCGACCGACCTTACGCTCCCTCCACCGCT	419					

Db	1028	GCAGATCTGAGAGTGGATCTCAACGACAGCGACCGCACTACGCTCTCTCCACCCAGGT	1087		
Qy	420	TTGATGATTATAAGACTCAAGCGCAACACAGCTTACGATACTATCTTTTACCTCAACATCAC	479		
Db	1088	TTGATGATTATAAGACTCAAGCGCAACACAGCTTACGATACTATCTTTTACCTCAACATCAC	1147		
Qy	480	TAGCTGATACATACAGGCTGCTTTGGTAGGCTTCAGAGTGTGTGTCATCAATAATAAGGATA	539		
Db	1148	TAGCTGATACATACAGGCTGCTTTGGTAGGCTTCAGAGTGTGTGTCATCAATAATAAGGATA	1207		
Qy	540	CAGCGGCTACTGATAGGAAACGCAATCGCTCGGAGTGGGAAATAAGAAATGCGCGATG	599		
Db	1208	CAGCGGCTACTGATAGGAAACGCAATCGCTCGGAGTGGGAAATAAGAAATGCGCGATG	1267		
Qy	600	CAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGA	659		
Db	1268	CAGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGA	1327		
Qy	660	TTCTTGACTCTTTAGGTAAACCTGACTCTCTTCGACCTCTTACAGGCTGCTTCTTCTCAAT	719		
Db	1328	TTCTTGACTCTTTAGGTAAACCTGACTCTCTTCGACCTCTTACAGGCTGCTTCTTCTCAAT	1387		
Qy	720	CTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAG	779		
Db	1388	CTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAG	1447		
Qy	780	TCCAGGGAACCGCTTGCATTTGCTCAATCTTTTAGTTGATCAGACAGATGCTTACAGCGA	839		
Db	1448	TCCAGGGAACCGCTTGCATTTGCTCAATCTTTTAGTTGATCAGACAGATGCTTACAGCGA	1507		
Qy	840	CACAGTAGAAGAAGATGGAAATGCGATAGGAGTGCATATTTTGCAGGACAGACGCTA	899		
Db	1508	CACAGTAGAAGAAGATGGAAATGCGATAGGAGTGCATATTTTGCAGGACAGACGCTA	1567		
Qy	900	GTGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAG	959		
Db	1568	GTGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAGCTAAAG	1627		
Qy	960	CAGCAATCGCTACTGTCTAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTC	1019		
Db	1628	CAGCAATCGCTACTGTCTAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTC	1687		
Qy	1020	CAATTTCTCAGAGCGGACAAATGGTATATACAGCTGGAAGATCTTTAAATAATCA	1079		
Db	1688	CAATTTCTCAGAGCGGACAAATGGTATATACAGCTGGAAGATCTTTAAATAATCA	1747		
Qy	1080	AACCTGCAGATGGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAAC	1139		
Db	1748	AACCTGCAGATGGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAAC	1807		
Qy	1140	AAGGAAGTAGTATTTGATGATTTTCGTGTTTCCATGCTGTTAGATGATGCTGAAATGAGA	1199		
Db	1808	AAGGAAGTAGTATTTGATGATTTTCGTGTTTCCATGCTGTTAGATGATGCTGAAATGAGA	1867		
Qy	1200	CCGCTTCCATTTTGTGATGCTGGGTTTCGTGATGATTTTCCATGCTGTTAGATGATGCTG	1259		
Db	1868	CCGCTTCCATTTTGTGATGCTGGGTTTCGTGATGATTTTCCATGCTGTTAGATGATGCTG	1927		
Qy	1260	CTGATTTCTCAAGCTGCCCAACAGAGCTCGACACAAGCTAGACGAGGAAAGCGCGCTG	1319		
Db	1928	CTGATTTCTCAAGCTGCCCAACAGAGCTCGACACAAGCTAGACGAGGAAAGCGCGCTG	1987		
Qy	1320	GAGATCACAGTGTGCTGCGAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGGGCTCTAG	1379		
Db	1988	GAGATCACAGTGTGCTGCGAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGGGCTCTAG	2047		
Qy	1380	GTAAGCTGGCAACAAACAGGGCATCTCAATCTTTTAGACAGATCGCTTCTGCTGCTG	1439		
Db	2048	GTAAGCTGGCAACAAACAGGGCATCTCAATCTTTTAGACAGATCGCTTCTGCTGCTG	2107		
Qy	1440	TTGTGAGCGCAGGAGT	1455		
Db	2108	TTGTGAGCGCAGGAGT	2123		

RESULT 9
AR310754.08/c
WPCOMMENT
Sequence split into 13 fragments LOCUS AR310754 Accession AR310754

Fragment Name	Begin	End
AR310754_00	1	110000
AR310754_01	100001	210000
AR310754_02	200001	310000
AR310754_03	300001	410000
AR310754_04	400001	510000
AR310754_05	500001	610000
AR310754_06	600001	710000
AR310754_07	700001	810000
AR310754_08	800001	910000
AR310754_09	900001	1010000
AR310754_10	1000001	1110000
AR310754_11	1100001	1210000
AR310754_12	1200001	1230025

Continuation (9 of 13) of AR310754 from base 800001 (AR310754 Sequence 1 from patent US 99.2%; Score 1444; DB 6; Length 110000; Best Local Similarity 99.9%; Pred. No. 5.5e-299; Matches 1455; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	1	ATATAATCTTTAAACAGCGCTCGCATTAATTAATTAGTGAGAGCTTTTTTTTTTTTTTTT 60
Db	28698	ATATAATCTTTAAACAGCGCTCGCATTAATTAATTAGTGAGAGCTTTTTTTTTTTTTTTT 28639
QY	61	ATAATAAACTAAAGATTTTATTTATTTTATGATTTTATGTTTAAATCCATTGGTCC 120
Db	28638	ATAATAAACTAAAGATTTTATTTATTTTATGATTTTATGTTTAAATCCATTGGTCC 28579
QY	121	AGTCTCTATAGACGAAACAGACACCTCCGAGATCTTTCTGCTCAAGGATTGGA 180
Db	28578	AGTCTCTATAGACGAAACAGACACCTCCGAGATCTTTCTGCTCAAGGATTGGA 28519
QY	181	GGCGAGTGAGCAATTAAGAGTGGGAGCTCAAGAAATAGCAGTGGCGGAGCTTAAGCC 240
Db	28518	GGCGAGTGAGCAATTAAGAGTGGGAGCTCAAGAAATAGCAGTGGCGGAGCTTAAGCC 28459
QY	241	TAAAGATCTTAAGACCGATCTGTAGAGGATGGAGCATCTTTCGCTTCAGTGAATGC 300
Db	28458	TAAAGATCTTAAGACCGATCTGTAGAGGATGGAGCATCTTTCGCTTCAGTGAATGC 28399
QY	301	TCTATGAGCTGGGAGATTAAGTGGGATTTGCTTCTAGTAAACAGCTCGCTTCCTACTAG 360
Db	28398	TCTATGAGCTGGGAGATTAAGTGGGATTTGCTTCTAGTAAACAGCTCGCTTCCTACTAG 28339
QY	361	CAGATCTGAGACGTGGATCTAACGACGCGACCGACCTAGCCTTCCTCCACCGATTT 420
Db	28338	CAGATCTGAGACGTGGATCTAACGACGCGACCGACCTAGCCTTCCTCCACCGATTT 28279
QY	421	TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 480
Db	28278	TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 28219
QY	481	AGTGCATACAGCTGCTTTTGTGAGCTCCAGGATGCTGTCACTAATATAAGGATAC 540
Db	28218	AGTGCATACAGCTGCTTTTGTGAGCTCCAGGATGCTGTCACTAATATAAGGATAC 28159
QY	541	ACGGGCTACTGATGAGAAACCGCAATCGCTGGAGTGGAACTTAAGATGCCGATGC 600
Db	28158	ACGGGCTACTGATGAGAAACCGCAATCGCTGGAGTGGAACTTAAGATGCCGATGC 28099
QY	601	AGTTAAAGTTGGCGCAATTAACAGATTAAGCAATATGCTTCGGATAACCAAGCGAT 660
Db	28098	AGTTAAAGTTGGCGCAATTAACAGATTAAGCAATATGCTTCGGATAACCAAGCGAT 28039
QY	661	TCTTGACTCTTAGGTAAACAGTACTCTTCCAGCTCTTACAGGCTGCTTCTTCCAAATC 720
Db	28038	TCTTGACTCTTAGGTAAACAGTACTCTTCCAGCTCTTACAGGCTGCTTCTTCCAAATC 27979

QY	721	TGTAGCAAAACAATAACAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db	27978 <th>TGTAGCAAAACAATAACAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGT 27919</th>	TGTAGCAAAACAATAACAAGCAGCTGAGCTTCTTTAAAGAGATGCAAGATAACCCAGTAGT 27919
QY <th>781</th> <th>CCCAGGGAAACGCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840</th>	781	CCCAGGGAAACGCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db	27918	CCCAGGGAAACGCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 27859
QY <th>841</th> <th>ACAGATAGAGAAAGATGGAAATGCGATTAGGGATGCAATATTTTTCAGGACAGAAACGCTAG 900</th>	841	ACAGATAGAGAAAGATGGAAATGCGATTAGGGATGCAATATTTTTCAGGACAGAAACGCTAG 900
Db	27858	ACAGATAGAGAAAGATGGAAATGCGATTAGGGATGCAATATTTTTCAGGACAGAAACGCTAG 27799
QY <th>901</th> <th>TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGTTTTCAGTAAAGC 960</th>	901	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGTTTTCAGTAAAGC 960
Db	27798	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGTTTTCAGTAAAGC 27739
QY <th>961</th> <th>AGCAATCGCTACTGCTAAGACACAATAAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020</th>	961	AGCAATCGCTACTGCTAAGACACAATAAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
Db	27738	AGCAATCGCTACTGCTAAGACACAATAAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 27679
QY <th>1021</th> <th>AATTCTTCAAGAGCGGAAACAAATGGTAATACAGCTGAGAAAAGATCTTAAAAATATCAA 1080</th>	1021	AATTCTTCAAGAGCGGAAACAAATGGTAATACAGCTGAGAAAAGATCTTAAAAATATCAA 1080
Db	27678	AATTCTTCAAGAGCGGAAACAAATGGTAATACAGCTGAGAAAAGATCTTAAAAATATCAA 27619
QY <th>1081</th> <th>ACCTGCAGATGGTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGCAACA 1140</th>	1081	ACCTGCAGATGGTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGCAACA 1140
Db	27618	ACCTGCAGATGGTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGCAACA 27559
QY <th>1141</th> <th>AGGAAGTAGTATGCTAGTATTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200</th>	1141	AGGAAGTAGTATGCTAGTATTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
Db	27558	AGGAAGTAGTATGCTAGTATTCGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 27499
QY <th>1201</th> <th>CGCTTCCATTTTGTGCTGGTTTTCGTCAGATGATTCACATGTTTCAATACGAAAATCC 1260</th>	1201	CGCTTCCATTTTGTGCTGGTTTTCGTCAGATGATTCACATGTTTCAATACGAAAATCC 1260
Db	27498	CGCTTCCATTTTGTGCTGGTTTTCGTCAGATGATTCACATGTTTCAATACGAAAATCC 27439
QY <th>1261</th> <th>TGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACAGCGAAAGCGCTGG 1320</th>	1261	TGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACAGCGAAAGCGCTGG 1320
Db	27438	TGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACAGCGAAAGCGCTGG 27379
QY <th>1321</th> <th>AGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGGCTCTAGG 1380</th>	1321	AGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGGCTCTAGG 1380
Db	27378	AGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAGCGGCTCTAGG 27319
QY <th>1381</th> <th>TAAAGCTGGGCAACAAAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGTG- 1439</th>	1381	TAAAGCTGGGCAACAAAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGTG- 1439
Db	27318	TAAAGCTGGGCAACAAAGGCGATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGTG- 27259
QY <th>1440</th> <th>TTGTGAGCGCAGGAGT 1455</th>	1440	TTGTGAGCGCAGGAGT 1455
Db <th>27258</th> <th>TTGTGAGCGCAGGAGT 27243</th>	27258	TTGTGAGCGCAGGAGT 27243

RESULT 10
AX338293
LOCUS
DEFINITION Sequence 63 from Patent WO0181379.
ACCESSION AX338293
VERSION AX338293.1 GI:18128828
KEYWORDS
SOURCE Chlamydophila pneumoniae
ORGANISM Chlamydophila pneumoniae
REFERENCE Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomphila.
AUTHORS Bhatia, A., Probst, P. and Stromberg, E.J.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0181379-A 63 01-NOV-2001;
CORIXA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..1956
/organism="Chlamydomphila pneumoniae"

ORIGIN		/mol_type="unassigned DNA" /db_xref="taxon:83558"		93.1%; Score 1355; DB 6; Length 1956; Best Local Similarity 100.0%; Pred. No. 1e-279; Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	101	ATGGTTAATCCTATTTGGTCAGGTCCTATATAGACGAACAGACGACACCTCCCGCAGAT	160		
Db	1	ATGGTTAATCCTATTTGGTCAGGTCCTATATAGACGAACAGACGACACCTCCCGCAGAT	60		
QY	161	CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAATAAGAGTCGGGAAGCTCAAGAAATA	220		
Db	61	CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAATAAGAGTCGGGAAGCTCAAGAAATA	120		
QY	221	GCAGGTGCGGAAGCTAAGCCCTAAAGAACTTAAGACCCGATTCCTGTAGAGCGATGGAGCATC	280		
Db	121	GCAGGTGCGGAAGCTAAGCCCTAAGAACTTAAGACCCGATTCCTGTAGAGCGATGGAGCATC	180		
QY	281	TTGGGTTCTGAGTAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTCTAGT	340		
Db	181	TTGGGTTCTGAGTAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTCTAGT	240		
QY	341	AACAGCTCGTCTTCTACTAGCAGATCTGCAGAGCTGGAGTCAACGACGACGACCGACCT	400		
Db	241	AACAGCTCGTCTTCTACTAGCAGATCTGCAGAGCTGGAGTCAACGACGACGACCGACCT	300		
QY	401	ACGCTCTCTCCACCCAGCTTTGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATPACT	460		
Db	301	ACGCTCTCTCCACCCAGCTTTGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATPACT	360		
QY	461	ATCTTTACTTCAACATCAGTACGATGACATACAGCTGCTGTTGGTAGCCCTCCAGGATGCT	520		
Db	361	ATCTTTACTTCAACATCAGTACGATGACATACAGCTGCTGTTGGTAGCCCTCCAGGATGCT	420		
QY	521	GTCACTAATAATAAGGATACAGCGGCTACTGATGAGGAACCGCAACTCGCTGGGAGTGG	580		
Db	421	GTCACTAATAATAAGGATACAGCGGCTACTGATGAGGAACCGCAACTCGCTGGGAGTGG	480		
QY	581	GAACAATAAGATGCGGATGAGTTAAAGTTGGGCGCAAAATTACAGAAATTAGCGAAATAT	640		
Db	481	GAACAATAAGATGCGGATGAGTTAAAGTTGGGCGCAAAATTACAGAAATTAGCGAAATAT	540		
QY	641	GCCTCGATTAACAGGATCTTGTACTCTTAGGTAACATGACTTCTCTCGACCTCTTA	700		
Db	541	GCCTCGATTAACAGGATCTTGTACTCTTAGGTAACATGACTTCTCTCGACCTCTTA	600		
QY	701	CAGGCTGCTCTTCTCCAACTGTAGCAAAACAATAACAGCAGCTGAGCTTCTTAAAGAG	760		
Db	601	CAGGCTGCTCTTCTCCAACTGTAGCAAAACAATAACAGCAGCTGAGCTTCTTAAAGAG	660		
QY	761	ATGCAAGATAACCCAGTAGTCCCAGGGAACCGCTGCAATTGCTCAATCTTTAGTTGAT	820		
Db	661	ATGCAAGATAACCCAGTAGTCCCAGGGAACCGCTGCAATTGCTCAATCTTTAGTTGAT	720		
QY	821	CAGACATGCTACAGCGACACATAGAGAAAGATGGAATGCCATAGGATGCATAT	880		
Db	721	CAGACATGCTACAGCGACACATAGAGAAAGATGGAATGCCATAGGATGCATAT	780		
QY	881	TTTGCAGGACAGAACCGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC	940		
Db	781	TTTGCAGGACAGAACCGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC	840		
QY	941	AACATAGATTAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG	1000		
Db	841	AACATAGATTAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG	900		
QY	1001	AAAAAGTCCCGGACTCTCCAAATTTCTCAAGACGGGAACAAATGGTAAATACAGGCTGAG	1060		
Db	901	AAAAAGTCCCGGACTCTCCAAATTTCTCAAGACGGGAACAAATGGTAAATACAGGCTGAG	960		
QY	1061	AAAGATCTTAAAAATATCAAACTGCAGATGGTCTGTGATGTTCCAAATCCAGGAACCTACA	1120		
RESULT 11					
AX349491					
LOCUS AX349491 1956 bp DNA linear PAT 06-FEB-2002					
DEFINITION Sequence 14 from Patent WO0202606.					
ACCESSION AX349491					
VERSION AX349491.1 GI:18615352					
KEYWORDS Chlamydia pneumoniae					
SOURCE Chlamydia pneumoniae					
ORGANISM Chlamydia pneumoniae					
REFERENCE 1					
AUTHORS Ratti, G. and Grandi, G.					
TITLE Immunisation against Chlamydia pneumoniae					
JOURNAL Patent: WO 0202606-A 14 10-JAN-2002;					
Chiron S.p.A. (IT)					
FEATURES					
source					
1..1956					
/organism="Chlamydia pneumoniae"					
/mol_type="unassigned DNA"					
/db_xref="taxon:83558"					
ORIGIN					
Query Match 93.1%; Score 1355; DB 6; Length 1956;					
Best Local Similarity 100.0%; Pred. No. 1e-279;					
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	101	ATGGTTAATCCTATTTGGTCAGGTCCTATATAGACGAACAGACGACACCTCCCGCAGAT	160		
Db	1	ATGGTTAATCCTATTTGGTCAGGTCCTATATAGACGAACAGACGACACCTCCCGCAGAT	60		
QY	161	CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAATAAGAGTCGGGAAGCTCAAGAAATA	220		
Db	61	CTTTCTGCTCAAGATTGGAGGCGAGTCGACGAATAAGAGTCGGGAAGCTCAAGAAATA	120		
QY	221	GCAGGTGCGGAAGCTAAGCCCTAAGAACTTAAGACCCGATTCCTGTAGAGCGATGGAGCATC	280		
Db	121	GCAGGTGCGGAAGCTAAGCCCTAAGAACTTAAGACCCGATTCCTGTAGAGCGATGGAGCATC	180		
QY	281	TTGGGTTCTGAGTAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTCTAGT	340		
Db	181	TTGGGTTCTGAGTAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTTCTAGT	240		
QY	341	AACAGCTCGTCTTCTACTAGCAGATCTGCAGAGCTGGAGTCAACGACGACGACCGACCT	400		
Db	241	AACAGCTCGTCTTCTACTAGCAGATCTGCAGAGCTGGAGTCAACGACGACGACCGACCT	300		

401 ACGCTCTCCACCCAGTTTGGATTAATAGACTCAGCGCAACAGCTTACGATCT 460
Db |||||
301 ACGCTCTCCACCCAGTTTGGATTAATAGACTCAGCGCAACAGCTTACGATCT 360
Qy |||||
461 ATCTTTACCTCAACATCACTAGTGCATACAGGCTGCTTTGGTGAAGCTCCAGGATGCT 520
Db |||||
361 ATCTTTACCTCAACATCACTAGTGCATACAGGCTGCTTTGGTGAAGCTCCAGGATGCT 420
Qy |||||
521 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 580
Db |||||
421 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 480
Qy |||||
581 GAACTTAAGATGCGGATCAGTTAAAGTTGGCGCAAAATACAGAAATACGAAATAT 640
Db |||||
481 GAACTTAAGATGCGGATCAGTTAAAGTTGGCGCAAAATACAGAAATACGAAATAT 540
Qy |||||
641 GCTCGGATTAAGCAAGCGATCTTGACTCTTTAGGTAAGTCAATCTTTAGTTGAT 700
Db |||||
541 GCTCGGATTAAGCAAGCGATCTTGACTCTTTAGGTAAGTCAATCTTTAGTTGAT 600
Qy |||||
701 CAGGCTGCTCTTCTCAATCTGTAGCAAAATACAGGAGCTGAGCTTCTTAAAGAG 760
Db |||||
601 CAGGCTGCTCTTCTCAATCTGTAGCAAAATACAGGAGCTGAGCTTCTTAAAGAG 660
Qy |||||
761 ATGCAAGATAACCCAGTACTGCTCCAGGAAACCGCTGCAATGCTCAATCTTTAGTTGAT 820
Db |||||
661 ATGCAAGATAACCCAGTACTGCTCCAGGAAACCGCTGCAATGCTCAATCTTTAGTTGAT 720
Qy |||||
821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATAGGAGTGCATAT 880
Db |||||
721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATAGGAGTGCATAT 780
Qy |||||
881 TTTCAGGACAGAACCTAGTGGAGCTGTAGAAATGCTAAATCTTAATACAGTATAAGC 940
Db |||||
781 TTTCAGGACAGAACCTAGTGGAGCTGTAGAAATGCTAAATCTTAATACAGTATAAGC 840
Qy |||||
941 AACATAGATTCAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAG 1000
Db |||||
841 AACATAGATTCAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTGAG 900
Qy |||||
1001 AAAAGTTCCCGACTCTCAATCTTCAAGAAAGCGGAAACAAATGTAATACAGGCTGAG 1060
Db |||||
901 AAAAGTTCCCGACTCTCAATCTTCAAGAAAGCGGAAACAAATGTAATACAGGCTGAG 960
Qy |||||
1061 AAGATCTTAAATATCAAACTGAGATGCTGCTGATGCTCCAAATCCAGGAACTACA 1120
Db |||||
961 AAGATCTTAAATATCAAACTGAGATGCTGCTGATGCTCCAAATCCAGGAACTACA 1020
Qy |||||
1121 GTTGGAGGCTCCAAGCAACAAAGAAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTA 1180
Db |||||
1021 GTTGGAGGCTCCAAGCAACAAAGAAAGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTA 1080
Qy |||||
1181 GATGATGCTGAAATGAGACCGCTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1240
Db |||||
1081 GATGATGCTGAAATGAGACCGCTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy |||||
1241 ATGTTCAATACGGAATCTGATCTCAAGCTGCCCAACAGAGGCTCGCAGACCAAGCT 1300
Db |||||
1141 ATGTTCAATACGGAATCTGATCTCAAGCTGCCCAACAGAGGCTCGCAGACCAAGCT 1200
Qy |||||
1301 AGAGCAGGAAACCGCTGAGATGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1360
Db |||||
1201 AGAGCAGGAAACCGCTGAGATGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy |||||
1361 GCTTTAGAGCGGCTCTAGTAAAGCTGGCAACAGGAGTACTCAATGCTTTAGGA 1420
Db |||||
1261 GCTTTAGAGCGGCTCTAGTAAAGCTGGCAACAGGAGTACTCAATGCTTTAGGA 1320
Qy |||||
1421 CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1485
Db |||||
1321 CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385

RESULT 12
AX361992 1956 bp DNA linear PAT 15-FEB-2002
LOCUS
DEFINITION Sequence 385 from Patent WO0208267.
ACCESSION AX361992
KEYWORDS AX361992.1 GI:18694447
SOURCE
ORGANISM Chlamydomophila pneumoniae
Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE 1
AUTHORS Fling, S.P., Skeiky, Y.A., Probst, P. and Bhatia, A.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0208267-A 385 31-JAN-2002;
CORIXA CORPORATION (US)
FEATURES
source
1. 1956
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
ORIGIN
Query Match 93.1%; Score 1355; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 1e-279;
Matches 1355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 ATGGTTAATCTTATGGTCCAGGTCCTATAGACGAAACAGACGACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTTATGGTCCAGGTCCTATAGACGAAACAGACGACACCTCCCGCAGAT 60
QY 161 CTTTCCTGCTCAAGGATTCGAGCGAGTCAGCAATATAGAGTGGGAGCTCAAAAGATA 220
Db 61 CTTTCCTGCTCAAGGATTCGAGCGAGTCAGCAATATAGAGTGGGAGCTCAAAAGATA 120
QY 221 GCAGTGGCGAAGCTAAGCCTAAAGAACTTAAGACCGATCTGTAGAGCGATGGAGATC 280
Db 121 GCAGTGGCGAAGCTAAGCCTAAAGAACTTAAGACCGATCTGTAGAGCGATGGAGATC 180
QY 281 TTGCGCTTCGAGTGAATGCTCTCATGATCTGGCAGATAGCTGGTATTGCTTCTAGT 340
Db 181 TTGCGCTTCGAGTGAATGCTCTCATGATCTGGCAGATAGCTGGTATTGCTTCTAGT 240
QY 341 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCCACT 400
Db 241 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCCACT 300
QY 401 AGCCTCTCCACCCAGCTTTGATGATTAAGACTCAAGCGCAACAGCTTACGATCT 460
Db 301 AGCCTCTCCACCCAGCTTTGATGATTAAGACTCAAGCGCAACAGCTTACGATCT 360
QY 461 ATCTTTACTCAACATCACTAGTGCATACAGGCTGCTTTGGTGAAGCTCCAGGATGCT 520
Db 361 ATCTTTACTCAACATCACTAGTGCATACAGGCTGCTTTGGTGAAGCTCCAGGATGCT 420
QY 521 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 580
Db 421 GTCACTAATATAAGGATACAGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTG 480
QY 581 GAACTTAAGATGCGGATCAGTTAAAGTTGGCGCAAAATACAGAAATACGAAATAT 640
Db 481 GAACTTAAGATGCGGATCAGTTAAAGTTGGCGCAAAATACAGAAATACGAAATAT 540
QY 641 GCTTCGGAATCAACAGCGATCTTGACTCTTTAGGTAAGTCAATCTTTAGTTGAT 700
Db 541 GCTTCGGAATCAACAGCGATCTTGACTCTTTAGGTAAGTCAATCTTTAGTTGAT 600
QY 701 CAGGCTGCTCTTCTCAATCTGTAGCAAAATACAGGAGCTGAGCTTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCAATCTGTAGCAAAATACAGGAGCTGAGCTTCTTAAAGAG 660
QY 761 ATGCAAGATAACCCAGTACTGCTCCAGGAAACCGCTGCAATGCTCAATCTTTAGTTGAT 820


```
Db      781  GCAGATGGTTCGATGTTCCAAATCCAGAACTACAGTTGGAGCTCCAGACCAAGGA 840
QY      1145  AGTAGTATTGGTAGTATTGGTGTTCATGCTGTAGATGATGCTGAAATGAGACCGCT 1204
Db      841  AGTAGTATTGGTAGTATTGGTGTTCATGCTGTAGATGATGCTGAAATGAGACCGCT 900
QY      1205  TCCATTTTCATGCTCTGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT 1264
Db      901  TCCATTTTCATGCTCTGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT 960
QY      1265  TCTCAAGCTGCCCAACAGGAGCTCGCAGACAAGCTAGACGACGCAAAAGCCGCTGAGAT 1324
Db      961  TCTCAAGCTGCCCAACAGGAGCTCGCAGACAAGCTAGACGACGCAAAAGCCGCTGAGAT 1020
QY      1325  GACAGTCTGCTCGCAGCGCTGGCAGATGCTCAGAAAGCTTGAAGCGGCTCTAGTAAA 1384
Db      1021  GACAGTCTGCTCGCAGCGCTGGCAGATGCTCAGAAAGCTTGAAGCGGCTCTAGTAAA 1080
QY      1385  GCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGTCTCTGCTGTGTG 1444
Db      1081  GCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGTCTCTGCTGTGTG 1140
QY      1445  AGCGCAGGAGT 1455
Db      1141  AGCGCAGGAGT 1151

RESULT 14
LOCUS      AE016994
DEFINITION Chlamydomophila caviae GPIC section 1 of 4 of the complete genome.
ACCESSION  AE016994
VERSION     AE016994.1
KEYWORDS   GI:29834116
SOURCE     Chlamydomophila caviae GPIC
ORGANISM   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE  1 (bases 1 to 304769)
AUTHORS    Read,T.D., Myers,G.S., Brunham,R.C., Nelson,W.C., Paulsen,I.T.,
            Heidelberg,J., Holtzapple,E., Khouri,H., Fedorova,N.B., Carty,H.A.,
            Umayam,L.A., Haft,D.H., Peterson,J., Beanan,M.J., White,O.,
            Salzberg,S.L., Hsia,R.C., McClarty,G., Rank,R.G., Bavoil,P.M. and
            Fraser,C.M.
TITLE       Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
            examining the role of niche-specific genes in the evolution of the
            Chlamydiaceae
JOURNAL    Nucleic Acids Res. 31 (8), 2134-2147 (2003)
MEDLINE    2269155
PUBMED     12682364
REFERENCE  2 (bases 1 to 304769)
AUTHORS    Read,T., Myers,G., Brunham,R., Nelson,W., Paulsen,I.,
            Heidelberg,J., Holtzapple,E., Khouri,H., Fedorova,N., Carty,H.,
            Umayam,L., Haft,D., Peterson,J., Beanan,M., White,O., Salzberg,S.,
            Hsia,R.-C., McClarty,G., Rank,R., Bavoil,P. and Fraser,C.
TITLE       Direct Submission
JOURNAL    Submitted (29-OCT-2002) The Institute for Genomic Research, 9712
            Medical Center Dr. Rockville, MD 20850, USA
FEATURES   Location/Qualifiers
            source
              1. 304769
                /organism="Chlamydomophila caviae GPIC"
                /mol_type="genomic DNA"
                /isolate="GPIC"
                /db_xref="taxon:227941"
            gene
              57..1052
                /gene="hemB"
                /locus_tag="CCA00001"
            CDS
              57..1052
                /gene="hemB"
                /locus_tag="CCA00001"
                /notes="similar to GB:L16784, SP:P33395, and PID:290384;
                identified by sequence similarity; putative"
                /codon_start=1
```

```
/transl_table=11
/product="delta-aminolevulinic acid dehydratase"
/protein_id="AAP04754.1"
/db_xref="GI:29834117"
/translation="MSSIALRRPRNRKTAIRDLVAETALLPQDFICPPFVKRGKN
ICEETESLPVGRWSIDLLKEIERLCSLGLKAVILFPVIPSIDLKDAYGSSYNPKNI
LCKSIYEYKAPDPLCVISDIALDPYTHGDGIMDQGEVLNDESVRIFGNIATLHAE
MGADIVASDMMDGRAHRAKLQDSQWNTNLIISYVKYASSLYNPRDALGSLQS
GKNRYQNPVKNVLELLECEDEGADMLMKPAGLYLDVHLHVRNSTIPLLAAYQ
VSGEYALTAADRGVGLNKEAMFYESLIAIKRAGADMTISYATPLILELVSSKF"
complement(1068..2480)
/gene="nqrA"
/locus_tag="CCA00002"
complement(1068..2480)
/gene="nqrA"
/locus_tag="CCA00002"
/notes="similar to SP:Q56586, GB:L16784, SP:P33395, and
PID:290384; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, subunit A"
/protein_id="AAP04755.1"
/db_xref="GI:29834118"
/translation="MKIATIRGLDLQSPKESGFLKRIDPALVSVDIRPYSALAK
LKVEQDDVISGSPATKPNPFGVITSPVSGIVKEIRRGKRSLLDVVIRKTPGONL
TEYSYDLKSLRAELLEIFKKEGLFALFKQRPDIPALPTOTPRDFINLADNRPTP
STEKQLAVFSRRESEGFYFNVGRVAKLFGLCPIHVTDLALPEKILKSTAHLHKI
TGYPGSGPSTHIIATPNEKIOVITLSFQEVLTIGHLFKGLNMQVVALAGSG
LKSSLRYVITTKGANFESLLPQDISDVSLSGDPLTGLCRSLPCGLMRDSTI
SVLPNFKROAFNLRGLNGKPTHTRTLSGLFKKHTYMDPTNLHGETRPIIDTEI
YDKVMLKIPVVPVLTIKAVITKDFELACMLGFLVCPDFALPTFDPSKTEMLTIKD
ALKDYAKETGLNLEVEDKE"
complement(2502..2936)
/locus_tag="CCA00003"
complement(2502..2936)
/locus_tag="CCA00003"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP04756.1"
/db_xref="GI:29834119"
/translation="MNSNSAKIIDSIKOILTLYNIDFPSPGSLSDSEVDYELM
EKTQKIOELDKRSODIILQOTGMTEQMEVANNPDNPFSEWLALENIRSSCDAYK
BTEALINVTKDLGSETITSSNNKTKTSSKKNKKKNKNIPL"
3048..5201
/gene="greA"
/locus_tag="CCA00004"
3048..5201
/gene="greA"
/locus_tag="CCA00004"
/notes="similar to GB:L10328, SP:P04846, GB:M12163,
PID:L46649, PID:290510, and PID:L790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="transcription elongation factor GreA"
/protein_id="AAP04757.1"
/db_xref="GI:29834120"
/translation="MDYLEKLQVLIDEQPSFFNLWBEYCFNDVVVGAELVQILEKV
KHSLAPLFGKIADTVLPMWIPDGKEDRVLQVLDVQNTNAGKPFYDAAIDYVNRK
YAGRENFEALRVGLRGREFQYSLRFDPLMLHSEGNFVPHSGWGVGEWVSFL
QOKVLIIEFGVWMAKDISFETAPKSLPLEKDHFLSRFGPDGDFGSAKHPSPVIE
MLKDLGPKTAKEIKDELVDLVIPEAEWNRWQATKTKKNICTITPKTIKDYRYN
PGDSLITQLETRLSQIEDTSKLVIEYQIFRDLHSELKPKENREIVIKALQTLUSVEG
NQSLEIQRDLLSLDFLGKSVSLDSEFLSPDEEIVSVNNISIVILQKRAFLMIKK
HSPVNVNFMKIFLSTSPSRELTFKVFERGDSYRKKEIKKLECAHPNMYDEVFT
WFFKLKSHDDGIFKPNDKTERLFLAALVFMYHVASTPOKELGKTYTIVLSORVL
AVRNMEGAPLYLKEILLSTKCSQSSDLSVLQSLAEVVHVDPKNTVTVVEDIIL
WTTPBSFTKMKNLQSLSGKGMVNDAKEIEDARVLGDRENSKPFALKKARLQBEI
VTLSEVFNARILTKDVIITDKVGVGCKVLEEDGNLITYSILGFPDADPDNYILSL
KSKLAQEMLDKSVGAAGVQGGKYKISRISQIWDE"
5353..6543
gene
```

```

/locus_tag="CCA00005"
5353.._8543
/locus_tag="CCA00005"
/notes="similar to SP:P04693, GB:L10328, SP:P04846,
GB:M12163, PID:146649, PID:290510, GB:U00096, and
PID:1790093; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="amino acid biosynthesis aminotransferase, class
1"
/protein_id="AAP04758.1"
/db_xref="GI:29834121"
/transl_table=11
/translation="MSFENQLPFRSPDSTLGLQKLFLEDEREKNLVIGVEDPSKA
YGGPSVKAQPLFEDENKGYLPISGLSPNOEMKLVGDVNNSVFVGAQALGCT
GALHLGAKLFVAYTSGKYVPEQWGNHVRIPQAGLEVILKPIYSSSEKLLFDEM
LSVLAAPNSVLVLQCCNPTGMDLNMWIRLAEVKNKHLQPLFFEDTSLYGFGRG
IKEDRRPEIFTEVGNPVFVAACASKNFSLYGERVGFVAVHSKITEDLDKISCLLEEK
ARGEYSPFRKAKIVTILSDASLKKEWLSLSDTIRSLGKTRARLVQAMNHHVGH
FDPILSQKGFPGFPGFVEQLFLRSEKGIYVTSRGARFNLGITDKNIDHVSQSPAEA
YQLS"
6839.._7861
/locus_tag="CCA00006"
6839.._7861
/locus_tag="CCA00006"
/notes="similar to GB:L10328, SP:P04846, GB:M12163,
PID:146649, PID:290510, GB:U00096, and PID:1790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP04759.1"
/db_xref="GI:29834122"
/translation="MAGSVARRHKNLYLVVLVLSVILCWSLPRFFYENIKQHFV
SLYRVFKRSVPSIDTIQVENVILKORAILERLQAEVLYNPPVFFSEQSLP
YFRKTIKRVIRYDPAHWSSQWNVNGKQIKQNSPVLKVLVLYDYVGEQOSRI
SLITDVGQSPVIAVGGIOQAWLVKDOIQGLTKHLERLSDSYILEKXKVDQLEEL
SSIQCSDDNTLLAGTSLSGGGLMKDETLVHGEFGCFDQKGLRVGDILITITGLD
GVFPGLFVAEITKVCQREGACSYKBAKSLAEPLTNLSVLILLPAMEFENDRPDI
FGLLWD"
complement(7821.._10958)
/gene="recB"
/locus_tag="CCA00007"
complement(7821.._10958)
/gene="recB"
/locus_tag="CCA00007"
/notes="identified by match to TIGR protein family HMM
TIGR01866"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease V, beta subunit"
/protein_id="AAP04760.1"
/db_xref="GI:29834123"
/translation="MKPDIENPQTSIRGKVFLEASAGTKFTIEQIVLRALLGSGV
SHVENLAVTFNATNESEKURIQNLKQALCQIKSVLEDPKSLPPYKDTSNVKLL
YQVNAALATDRAVATIHGFCNYVLQQHPKMQMTQKNAALTHSQAVFHHKKYLS
QDLWSVLPFQFVLAAKYNKSHKTAFLTDKLLSYTLQTFEHLPTQSTTLDALRS
WHNSIRSKIQDIPKEDRQILQHKDGFKKQFPFIDEDLRFVECLVASETSVOLFSF
SKIAETHPKRLTEYKSKAFSHLEETLSLYTEQFCNIELINLLOLVQVLYKNH
YTMLSPDESIALDEILLSRSEIIEILRRKFOLVILDFQDTDKQWNIIFSKLFA
NDAFSGSLFGDPQKSYIERNADLATYKAKSFPKSSQLHNNYRSTQAQMEAV
NILCFKFFQIPQSYEYHPLCPQSEHFENSHAPIHFFSYDDISDQAAWISHT
ASYLQATVHILGRMALIVSDSAQAFDLTHCSIPVAFSKNSIFHLPEVYLLTIAWL
EAILHVENKIORVLLSLFPHNLSLDTIEKKLYSTYFSLRSVVPDYGILAFYHF
MTLQCEALLKTPQDLSQEMERLCAYLDTVSSHQQLLYLOVSETRWERSLSFS
SYSETEILKTTIHASKGLEYDVVFCGLDKSKKNSASEWIREMYVACFRANKQLF
IPQNSNRNSALNTYKGVHVSYSILDLAQLTKGHPRLFSLSITPESRSTDVP
YKIIIPSTFTLISPAKQIFSFVSKLAFDNEVLDNPVAFVAFVLPGRKGTGII
HKILEINPFIKISPSIVIVFVKNTHLEGEYIIISQKLVTFSSPLSPATESFS
LQDICPSKIACBESFLSNQBLQWGLDFEYNDKYIIDMKTFSFLESLYSKE
NLLNVIKQNLDYQALITHAARFLQFDIVSDVMGFIIFIRGMDPEGNGLFCLPNQ
KISNPIITTRKCPAYH"
complement(10945.._14010)
/gene="recC"

```

```

/locus_tag="CCA00008"
complement(10945.._14010)
/gene="recC"
/locus_tag="CCA00008"
/notes="similar to GB:L25079, SP:P42822, and PID:529423;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="exodeoxyribonuclease V, gamma subunit"
/protein_id="AAP04761.1"
/db_xref="GI:29834124"
Query Match 14.0%; Score 204.2; DB 1; Length 304769;
Best Local Similarity 50.4%; Freq. No. 1.9e-33;
Matches 746; Conservative 0; Mismatches 633; Indels 102; Gaps 6;
QY 1 ATAAAAATCTTTAAAAAAGAGCTCGCATTAATATTAGTGAGAGCTTTTTTTTATTTT 60
DB 21625 ATAAAAATCTTTAAAAAAGAGCTCACATCTAATTAATAAGATGTGAGCTTTT 21684
QY 61 ATAATAAACTTAAAGATTTTATTTATTTTGTAGTTTTTATGTTAACTTATGGTCC 120
DB 21685 TTATAATTTAAGAAAAGATTTTATTAGTTTTTTTATTATTATGTTAATCTCTGT 21738
QY 121 AGGTCTTATAGCAAAACAGAACACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 180
DB 21739 CGGCCCTATAGATGAATCAAAAACATTGCTCTCGCAGACTTACTTCTTAGGTATGCA 21798
QY 181 GCGAGTGCAGCAAAATAAGAGTGGGAGCTCAAGAATAGCAGTGGGAAGCTTAAGCC 240
DB 21799 GCGAGCGCAGCAAAATCGTAGCTCAGAAGCTCAATCGATAACCGAATTCGAGCAATC 21858
QY 241 TAAAGATCTAAGAC-----CGATTCTGTAGAGGATGAGAGCATCTTTCGTTCTGAGT 294
DB 21859 CGGTTCATCGAGCGCTTCTGTGAAACTGTAGAGCATGAGTCTTTTGTGCTCTGCTCG 21918
QY 295 GAATGCTCTCATGATGCTGCGAGATAGTGGTATTGCTTCTAGTAAACAGCTCGTCTTC 354
DB 21919 GAAAGTTTACCAAGTCTTTTCGATAAGAT----- 21948
QY 355 TACTAGCAGATCTGCAGACGTGGAGCTCAACGACAGCAGCCAGCTACGCTCTCTCCACC 414
DB 21949 -----TTCTCTGTTCTTTTCAGGGAACGACTCTCTCA 21981
QY 415 CAGTTTGTATGATTAAGACTCAAGCGCAAAACAGCTTACGATACTATCTTTTACCTCAAC 474
DB 21982 AACTTTTGTATGAAGCTAAGACGCAAGCAGAGAGTGCAGAAAACCTGCGTGCAGAGTGCAC 22041
QY 475 ATCACTAGCTCACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACATAATAAA 534
DB 22042 TACTTATGATCAGTTCAGACCGCTTTTACAGCAGCTGCAAGATGCTGTGAAACAGATGGA 22101
QY 535 GGAATACAGCGGCTACTGATAGGAAAACCGCAATCGCTSCGAGTGGGAAACTAAGAATGC 594
DB 22102 GCAATTAGCTACTACTGATGAGAAAAGACTACAGTTGCTACATGGAACCGGCTCTTGA 22161
QY 595 CGATGCAGTTAAAGTTGGCGGCAANTTACAGATTAGCGAAATATGCTTCGGAATACCA 654
DB 22162 GGCACAGAAGAGTAGTCGCTGGATACACTTAACAGTTGGGTGCTATTCTTACAGAGAACA 22221
QY 655 AGCGATCTTCTGACTCTTTTAGTAAACTGACTTCCTTCCAGCTCTTACAGGCTGCTCTTCT 714
DB 22222 GAACTTCTTGAGGCAATAAGACGACCTCGTCTATGATCAGATTATGGAGCTGCGGG 22281
QY 715 CCAATCTGTAGCAAAACATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCC 774
DB 22282 ACAAGTAGAACCAATAAACAACCTGCTGAGGAGCTTAATTAACAGTTGAAGGA----- 22335
QY 775 AGTAGTCCAGGGAACCGCTGAAATGCTCAATCTTTTAGTTCATCAGACAGATGCTAC 834
DB 22336 -----AGCTGGGTTAGCTATCTCTGTGATAGTACCTTTGAGAAGCAAAATTACACCTC 22389
QY 835 AGCGACACAGATAGAGAAAGATGAAATGGATTAGGATGCAATATTTTTCAGGACAGAA 894

```

```

Db 22390 AGNACTCAGGTTACTGAATTAGCAGATGCTATATCGAAGCTTATGCTCGGGGAAAA 22449
Qy 895 CGCTAGTGGAGCTGTGAAAAATGCTAAATCTAATAACAGTATATAGCAACATAGATTTCAGC 954
Db 22450 CAGTACCGCGCTGTGGGCAAGCACAGGCAAAATAACAGCCCGCAATATAGAGCTTC 22509
Qy 955 TAAAGCAGCATCCTCTACTGTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGA 1014
Db 22510 CAAACAAACTATTGCAATGCACAAAAGTCATAGAAGACGCTCTTAAACTTGTCTCCAGA 22569
Qy 1015 CTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAAATACAGGCTCAGAAAGATCTTAAAA 1074
Db 22570 TTCTCCGATACTCAAGCTGTTGAAGAAGACACACAGGCGACAAAGATATCTCTAA 22629
Qy 1075 TATCAAACTCCAGATGGTTCTGTATGTTCACAAATCCAGGAAC-----TACAGT 1122
Db 22630 TGTAAACCTTAGTGGTGTAGTGTGCTATCGGTGGTCTCTGGAGCTCCTGGTAGTGT 22689
Qy 1123 TGGAGGCTCCAAGCAACAGGAAGTAGTATGTGTATGCTGTTTCCATGCTGTTAGA 1182
Db 22690 GGGGACTTCTCAAAATCGCGTGTCTACTTGGGGGAAGTTTCGGGTATCGATGTTATTGAC 22749
Qy 1183 TGATGCTGAAATAGAGACCGCTTCCATTTTGATGCTCTGGGTTTCGTACAGATGATTCACAT 1242
Db 22750 TGAATTTGATAATGAACCCGACGAGATCATATGCAAGGTTTCAGAAATATATGTCGATA 22809
Qy 1243 GTTCAATACGGAAATCTGATCTCAAGTGCACACAGGAGTCCAGCAACAGCTA- 1301
Db 22810 CTTCATGATCAAAACTCTGATTTTACAGCGCCTTTTGAAGAGATTATGAATCAAGTAAC 22869
Qy 1302 -----GAGCAGCGAAAGCGCTGGAGATGACAGTGTCTGTCAGCGCTGGCAGATGC 1353
Db 22870 CGACTTATACCGCAGATCAATCTCGAGATCGGAAGCTTACAGCAACACTCAAGAAAT 22929
Qy 1354 TCAGAAAGCTTTAGAGCGGCTTAGTAAAGCTGGGCAACACAGGGGCATCTCAATGC 1413
Db 22930 ACAACAAACCATACAGATGCCCTTCAAGGAGCTGCCGCTCAAGACGCGATGATCAATGC 22989
Qy 1414 TTTAGGACAGATCGCTTCTGCTGCTGTGTGAGCGCAGGAG 1454
Db 22990 TTTAGGAGCTATAACAACAGCAGCTTCAATTTCTACAGGAG 23030

```

```

RESULT 15
BD243971
LOCUS BD243971 1550 bp DNA linear PAT 17-JUL-2003
DEFINITION Chlamydia antigens and corresponding DNA fragments and uses
thereof.
ACCESSION BD243971
VERSION BD243971.1 GI:33053741
KEYWORDS JP 2002528081-A/1.
SOURCE Chlamydia pneumoniae
ORGANISM Chlamydia pneumoniae
REFERENCE 1 (bases 1 to 1550)
Murdin,A.D., Oomen,R.P. and Dunn,P.L.
Chlamydia antigens and corresponding DNA fragments and uses thereof
Patent: JP 2002528081-A 1 03-SEP-2002;
CONNAUGHT LABORATORIES LTD
OS Chlamydia pneumoniae
PN JP 2002528081-A/1
PD 03-SEP-2002
PF 28-OCT-1999 JP 2000578453
PR 28-OCT-1998 US 60/106037,20-SEP-1999 US 60/154658 PR
26-OCT-1999 US 09/427501
PI ANDREW DAVID MURDIN,RAYMOND PETER OOMEN,PAMELA LESLEY DUNN PC
C12N15/09,C12N15/09,A61K38/00,A61K39/118,A61K39/39, PC
A61P9/10.
PC A61P11/00,A61P11/06,C07K14/22,C07K14/295,C07K16/12,C12N1/15, PC
C12N1/19,
PC C12N1/21,C12N5/10,C12P21/02,C12Q1/68,G01N33/53,G01N33/566, PC
G01N33/569//

```

```

PC (C12N1/21,C12R1:01),(C12N15/09,C12R1:01),(C12P21/02,C12R1:01),
PC C12N15/00,
PC C12N15/00,C12N5/00,A61K37/02,(C12N15/00,C12R1:01) CC
Chlamydia antigens and corresponding DNA fragments and uses CC
thereof
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
source 1..1550
/organism="Chlamydia pneumoniae"
/mol_type="genomic DNA"
/db_xref="taxon:83558"

```

ORIGIN

```

Query Match 11.7%; Score 170; DB 6; Length 1550;
Best Local Similarity 100.0%; Pred. No. 7.3e-26;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATAAATCTTTAAACACAGCTCGCATTAATTATTAGTCGAGAGCTTTTATTTT 60
Db 1381 ATAAATCTTTAAACACAGCTCGCATTAATTATTAGTCGAGAGCTTTTATTTT 1440
Qy 61 ATAAATCTTTAAACACAGCTCGCATTAATTATTAGTCGAGAGCTTTTATTTT 120
Db 1441 ATAAATCTTTAAACACAGCTCGCATTAATTATTAGTCGAGAGCTTTTATTTT 1500
Qy 121 AGTCTCTATAGAGCAACAGAGCAGACCTCCCGCAGATCTTCTGCTC 170
Db 1501 AGTCTCTATAGAGCAACAGAGCAGACCTCCCGCAGATCTTCTGCTC 1550

```

Search completed: March 24, 2004, 01:57:16
Job time : 5663.52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:19 ; Search time 3817.47 Seconds

(without alignments)
11389.568 Million cell updates/sec

Title: US-10-608-559-5

Perfect score: 1456

Sequence: 1 ataaaatttttaaacagg.....ctgttgtgagcaggagta 1456

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estnu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	52.8	3.6	843	28	BZ643413
C 2	52	3.6	1201	9	AL541034
C 3	50	3.4	1294	13	BX446742
C 4	49	3.4	909	13	BX453105

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C 5	48.4	3.3	1191	14	CD500670
C 6	48.2	3.3	533	13	EX353273
C 7	47.8	3.3	510	28	BH714761
C 8	47	3.2	808	29	CG815585
C 9	46.8	3.2	1201	13	EX356412
C 10	46.6	3.2	1201	9	AL531906
C 11	46.6	3.2	1202	28	CC262481
C 12	46	3.2	427	13	EX403499
C 13	46	3.2	1041	9	AL575932
C 14	46	3.2	1201	9	AL558113
C 15	45.8	3.1	642	13	EX396222
C 16	45.8	3.1	1104	13	EX396222
C 17	45.8	3.1	1201	9	AL561244
C 18	45.6	3.1	967	28	BH146682
C 19	45.6	3.1	967	29	CNS07PKF
C 20	45.4	3.1	528	12	BH340195
C 21	45.4	3.1	724	12	BH344100
C 22	45.4	3.1	819	28	BZ391755
C 23	45.4	3.1	1201	9	AL524496
C 24	45.2	3.1	336	14	CB098017
C 25	45.2	3.1	1004	13	EX334499
C 26	45.2	3.1	1072	13	EX359378
C 27	45	3.1	1164	28	CC261365
C 28	44.8	3.1	1124	13	EX436282
C 29	44.8	3.1	1151	10	BB964820
C 30	44.8	3.1	1173	9	AL551058
C 31	44.8	3.1	1183	13	EX379850
C 32	44.6	3.1	259	14	CB722003
C 33	44.6	3.1	1537	12	BH195792
C 34	44.4	3.0	856	28	BH132709
C 35	44.4	3.0	976	13	EX388320
C 36	44.4	3.0	1096	13	EX381337
C 37	44.4	3.0	1154	28	BZ636556
C 38	44.4	3.0	1201	9	AL548036
C 39	44.2	3.0	518	13	EX355479
C 40	44.2	3.0	591	9	AU286627
C 41	44.2	3.0	1201	9	AL531430
C 42	44.2	3.0	1201	13	EX458906
C 43	44	3.0	1174	9	AL546393
C 44	44	3.0	1201	13	EX334085
C 45	44	3.0	1201	13	EX364759

ALIGNMENTS

RESULT 1
BZ643413/c
LOCUS BZ643413 843 bp DNA linear GSS 29-JAN-2003
DEFINITION OGAOW57TM ZM 0.7_1.5 KB Zea mays genomic clone ZMMEM011J17,
genomic survey sequence.
ACCESSION BZ643413
VERSION BZ643413.1 GI:28104915
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 843)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Reenick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGAOW57TC
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR

CD500670 CDR46-D01
EX353273 EX353273
BH714761 BOMGV1TR
CG815585 SOYFPH83TV
EX356412 EX356412
AL531906 AL531906
CC262481 CH261-167
EX403499 EX403499
AL575932 AL575932
AL558113 AL558113
EX396222 EX396222
EX396222 EX396222
EX707012 EX707012
AL561244 AL561244
BH146682 025_N_14-
AL621633 13 end 01
BZ340195 BZ340195
BZ344100 BZ344100
BZ391755 EINH26TR
AL524496 AL524496
CB098017 ku48b04.Y
EX334499 EX334499
EX359378 EX359378
CC261365 CH261-72H
EX436282 EX436282
BE964820 601658475
AL551058 AL551058
EX379850 EX379850
CB722003 jhm603F04
BH195792 602754186
BH132709 ENTNGS9TF
EX388320 EX388320
EX381337 EX381337
BZ636556 SP_Bac008
AL548036 AL548036
EX355479 EX355479
AU286627 AU286627
AL531430 AL531430
EX458906 EX458906
AL546393 AL546393
EX334085 EX334085
EX364759 EX364759

```

Class: sheared ends.
Location/Qualifiers
source
1. .843
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM85A011117"
/notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylated filtered genomic DNA library"

ORIGIN
Query Match 3.6%; Score 52.8; DB 28; Length 843;
Best Local Similarity 47.8%; Pred. No. 2.3; Mismatches 0; Gaps 0;
Matches 153; Conservative 0; Indels 0;

QY 286 TTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATGCTCTAGTAACAG 345
DB 742 TACTGCTACTGCTACTGCTGCTACTGCTACTGCTGCTACTGCTACTGCTACTGCTAC 683
QY 346 CTGCTCTTCTACTAGCAGATCTGCAGACGCTGGAAGCTCAACAGACGACCGACCTAGGCC 405
DB 682 TACTGCTACTACTACTGCTACTGCTGCTACTACTACTACTGCTACTGCTACTGCTGC 623
QY 406 TCTCCACACACGCTTGTGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTT 465
DB 622 TACTGCTGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGCTACTGC 563
QY 466 TACTCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTCTAC 525
DB 562 TGCTGCTACTGCTACTGCTGCTACTGCTGCTGCTACTGCTACTGCTACTGCTACTGC 503
QY 526 TAATATAAGGATACAGCGCTACTGATGAGGAACCGGAATCGCTGCGAGTGGGAAC 585
DB 502 TACTACTACTGCTACTACTGCTACTGCTGCTGCTGCTGCTGCTACTGCTACTGCTACT 443
QY 586 TAAGATGCGCGAGTCAGTTA 605
DB 442 TACTACTACTGCTGCTGCTA 423

RESULT 2
AL541034/c 1201 bp mRNA linear EST 12-MAY-2003
LOCUS AL541034 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005Y21
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION AL541034
VERSION AL541034.2 GI:30544815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12871720.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 951.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE005AAL1NP1&cluster=951.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005AAL1NP1.
Location/Qualifiers
1. .1201

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB011ZG07"
/tissue_type="NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005Y21"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 3.6%; Score 52; DB 9; Length 1201;
Best Local Similarity 42.0%; Pred. No. 3.1; Mismatches 0; Gaps 0;
Matches 63; Conservative 35; Indels 0;

QY 1 ATAAATCTTTAAACACGCTCGCATTAATATTAGTGAGAGCTTTTATTTTATTTT 60
DB 1160 AAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1101
QY 61 ATAAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
DB 1100 WAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1041
QY 121 AGGTCTCTATAGACAAACAGACGACACC 150
DB 1040 VMMMMMMVMVMGGMGRCMACACSCARRAC 1011

RESULT 3
BX446742/c 1201 bp mRNA linear EST 22-MAY-2003
LOCUS BX446742 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION CL0BB011ZG07 5-PRIME, mRNA sequence.
ACCESSION BX446742
VERSION BX446742.1 GI:31025761
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8247.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BB011ZG07RP1&cluster=8247.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BB011ZG07RP1.
Location/Qualifiers
1. .1201

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB011ZG07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```


Query Match	3.3%;	Score	48.4;	DB	14;	Length	1191;
Best Local Similarity	48.5%;	Pred.	No. 17;				

Query Match	3.3%	Score 48.4;	DB 14;	Length 1191;
Best Local Similarity	48.5%;	Pred. No. 17;		
Matches 133;	Conservative	0;	Mismatches 141;	Indels 0;
				Gaps 0;

Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
The Center of Excellence in Soybean Research, Teaching and
Outreach, Southern Illinois University at Carbondale and Plant
Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 1797
Email: Ga4082@siu.edu, jlschultz@siu.edu, cdtown@igr.org (URL:
<http://bioinformatics.siu.edu>)
Clones approximating a minimum tiling path were re-arrayed from the
library master plates prior to sequencing.
For purposes of clone identification each clone name is a
concatenation of the original clone location and its new location
in the re-arrayed sequencing plates.
Seq primer: **GTAATACGACTCTACTATAGGC**
Class: BAC ends.

```

FEATURES
source
      1. .808
      location/Qualifiers
      /organism="Glycine max"
      /mol_type="genomic DNA"
      /cultivar="Forrest"
      /db_xref="taxon:3847"
      /clone="H46C04:SQ3N22"
      /clone_lib="LargeInsertSoybeanGenLib"
      /note="Organ: Leaves; Vector: pCL004541 (pBELOBAC11 EcoRI
      clones); Site:1: BstXI; Soybean (Glycine max (L.) Merr.)
      cv. Forrest seeds were grown in greenhouse for fourteen
      days. Nuclei were isolated and embedded in agarose,
      restriction digested with Hind III BamHI or EcoRI, large
      size DNA fragments were ligated in vector V41 (pCL004541)
      and electro transformed in DH10a cells. About 90,000
      clones from BAC libraries were fingerprinted with HindIII
      and Hae III. Version 2 (automatic build) Contigs were
      built from 78,001 fingerprints. Contigs were manually
      examined to find the best non redundant tile path through
      the contigs representing about 13,000 clones. The clones
      were end sequenced"

```

ORIGIN		were end sequenced.	
Query Match	3.2%;	Score 47;	DB 29; Length 808;
Best Local Similarity	64.0%;	Pred. No. 34;	
Matches	71;	Conservative 0;	Mismatches 40; Indels 0; Gaps 0
QY	3	AAAATCTTTAAAAACAGCGCTCGCAATTAATTATAGTGAGAGCGCTTTTTTTTATTTTTTAT	62
Db	625	AAACTTTCAGAGAAAAACGTCGATTGATTTCTTGATTATTTTTTATTTTTTATTTTTT	566
QY	63	AATAAACTAAACAGATTTTTTATTTATTTTTTTCGAGTTTTTATCGTTAATCCTA	113
Db	565	CAAGATATTTTGATTTATTTTATTTATTTATTTTTCGCTTTTTTTCGTTTAAACCGA	515

RESULT	9
EX356412/c	
LOCUS	
DEFINITION	1201 bp mRNA linear EST 05-MAY-2003 BX356412 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI011YJ12 5-PRIME, mRNA sequence.
ACCESSION	BX356412
VERSION	BX356412.1
KEYWORDS	GI:30382071
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1201) Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of

```

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSODI011DE06QF1.
FEATURES
    Location/Qualifiers
        1..1201
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CSODI011YJ12"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-cclgga(dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

ORIGIN
Query Match          3.2%; Score 46.8; DB 13; Length 1201;
Best Local Similarity 27.2%; Pred. No. 35;
Matches 66; Conservative 75; Mismatches 102; Indels 0; Gaps 0

QY      55 TTTTATAAATAAACTAAAAGATTTTTATTATTTTGAGTTTTTATGTTTAATCCTAT 114
       |||:|||||:|||:|||:|||:|||:|||:|||:|||:|||
Db      914 TWTTYTTWAAAAATTTTWATTTTWWAATTTTTTYCAATYYTAWTTTAMCCCAA 855

QY      115 TGTCAGGTCTTAGACGAACAAGCACACCTCCCAGCATCTTCTGCTCAAGG 174
       |||:|||::|||::|||::|||::|||::|||::|||::|||
Db      854 AAATYTYMAAAAAAAATAAAMMMMAAANYCMGCMHTTTADKITTCCKKXBCMGGK 795

QY      175 ATTGGAGGCAGGTGCAGCAATAAGACTGCCGAGCTCAAGAATAGCAGGTGCGGAGC 234
       |||::|||::|||::|||::|||::|||::|||::|||::|||
Db      794 KKKKKKKKKKKKKKKDAAAAAAAAAADAAACCCTCCMMAAAAKKKKAABKBWAAA 735

QY      235 TAAGCCTTAAAGAACTTAAGACCGATTTCTGTAGAGCGATGGAGCATTTTGCCTTCTGCAAGT 294
       |||:|||::|||::|||::|||::|||::|||::|||::|||
Db      734 ADACAACAANAANCKKKKKKMSKKKKXBCBKAAAKKKBCCKKKKKKKKKKAANK 675

QY      295 GAA 297
       |||
Db      674 AAA 672

```

RESULT 10	AL531906	1201 bp	linear	EST 23-MAY-2003
LOCUS	AL531906			
DEFINITION	AL531906 Homo sapiens FETAL LIVER Homo sapiens cDNA clone			
	CS0DM003VL08 3-PRIME, mRNA sequence.			
ACCESSION	AL531906			
VERSION	AL531906.2	GI:31069738		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1201)			
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished (2001)			
COMMENT	On Feb 13, 2001 this sequence version replaced gi:127953199.			

FEATURES
source

```

/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match          3.2%; Score 46.6; DB 9; Length 1201;
Best Local Similarity 48.2%; Pred. No. 38;
Matches 67; Conservative 19; Mismatches 53; Indels 0; Gaps 0;

QY 28 TAATTATTAGTGGAGCTTTTATTTTATTTTATTAATAAACTAAAGATTTTATTTAT 87
Db 670 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 723
QY 88 TTTTGTAGTTTATGTTTATCTCTATTGTCAGGTCCTATAGACGAAACAGACGCAC 147
Db 730 TTTTATTTATTAAGATGAAAGAAAGAAAGACAGCTTTATATATATATATAAATAA 789
QY 148 ACCTCCGCGAGACTTTCT 166
Db 790 ATWTAATAATACATTTTAT 808

RESULT 11
LOCUS CC262481
DEFINITION CH261-167M9 Sp6.1 CH261 Gallus gallus genomic clone CH261-167M9,
genomic survey sequence.
ACCESSION CC262481
VERSION CC262481.1 GI:30607397
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 1202)
AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 30
High quality sequence stop: 105.
FEATURES
Location/Qualifiers
1..1202
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-167M9"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/notes="Vector: pRABAC2.1; Site1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match          3.2%; Score 46.6; DB 28; Length 1202;
Best Local Similarity 58.2%; Pred. No. 38;
Matches 82; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match          3.2%; Score 46; DB 13; Length 427;
Best Local Similarity 31.5%; Pred. No. 59;
Matches 34; Conservative 47; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTATTTT 60
Db 1004 ATAAATCTTTTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1063
QY 61 ATAAATAAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 1064 ATAAATAAATATAAATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAA 1123
QY 121 AGGTCTCTATAGACGAAACAGA 141
Db 1124 ATTATATTATAATAAAAAATA 1144

RESULT 12
LOCUS BX403499
DEFINITION BX403499 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA002H05
3-PRIME, mRNA sequence.
ACCESSION BX403499
VERSION BX403499.1 GI:30635017
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 427)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA002H05FP1.
FEATURES
Location/Qualifiers
1..427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA002H05"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match          3.2%; Score 46; DB 13; Length 427;
Best Local Similarity 31.5%; Pred. No. 59;
Matches 34; Conservative 47; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTATTTT 60
Db 64 WWWWWWWWWWWTWWTWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 123
QY 61 ATAAATAAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 108
Db 124 WTWWTAAATWAAATWATWATWATTTTATTTTATTTTATTTTATTTTATTTTAA 171

RESULT 13
LOCUS AL575932/c
DEFINITION AL575932 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1067YN05 3-PRIME, mRNA sequence.
ACCESSION AL575932
VERSION AL575932.2 GI:31314228
KEYWORDS EST.

```


/note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dT primed from Suq of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match	3.1%;	Score 45.8;	DB 13;	Length 642;
Best Local Similarity	50.7%;	Pred. No. 61;		
Matches 110;	Conservative 0;	Mismatches 107;	Indels 0;	Gaps 0;

QY	45	TTTTTTTTTTTATATAAATAAACTAAAGATTTTATTTTGTAGTTTTTATGG	104
Db	640		
QY	105	TTTTCTCTTTTATATAAATAAAACATAGCGAGACTTTCTCTTTCTTTTACAA	581
Db	580		
QY	165	TTAATCCTATTGTCCAGGTCCTATAGACGAACAGACACACCTCCCGCAGATCTTT	164
Db	520		
QY	225	GTGCGAAGCTAAGCCTAAGAACTCTAAGACCGATTC	261
Db	460		

Search completed: March 24, 2004, 05:56:12
Job time : 3820.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:04 ; Search time 830.202 Seconds
(without alignments)
11032.393 Million cell updates/sec

Title: US-10-608-559-1

Perfect score: 2156

Sequence: 1 ataaatttttaaaacagg.....gatcttcaataataaaaga 2156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002as.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2156	100.0	2156	3 AAD02063	Aad02063 Chlamydia
2	2143.4	99.4	110000	2 AAX91990_08	Continuation (9 of
3	1956	90.7	1956	4 AAS57031	Aas57031 C. pneumo
4	1956	90.7	1956	6 ABL92612	Ab192612 Chlamydia
5	1956	90.7	1956	6 ABL91190	Ab191190 Chlamydia
6	1956	90.7	1956	9 ADD42825	Add42825 Chlamydia
7	1848.8	85.8	1852	3 AAD02064	Aad02064 5'-trunca
8	1465.8	68.0	2238	3 AAD02066	Aad02066 C. pneumo
9	1455	67.5	1456	3 AAD02065	Aad02065 3'-trunca
10	187.6	8.7	110000	2 AAZ01425_07	Continuation (8 of
11	187.6	8.7	110000	2 AAZ01425_08	Continuation (9 of
12	187.2	8.7	1537	3 AAB63292	Aab63292 C. tracho
13	187.2	8.7	1537	4 AAB56196	Ab192425 Chlamydia
14	187.2	8.7	1537	6 ABL92425	Ab192425 Chlamydia
15	186	8.6	1944	9 ADD43867	Add43867 Chlamydia
16	185.4	8.6	1941	6 ABL92619	Ab192619 Chlamydia
17	184	8.5	1171	4 AAS56981	Aas56981 C. tracho
18	184	8.5	1171	9 ADD42775	Add42775 Chlamydia
19	184	8.5	1834	4 AAS57006	Aas57006 C. tracho
20	184	8.5	1834	9 ADD42800	Add42800 Chlamydia
21	183	8.5	1983	4 AAS56996	Aas56996 C. tracho
22	183	8.5	1983	9 ADD42790	Add42790 Chlamydia
23	170	7.9	1550	3 AAZ01425_07	Aaz01425 Chlamydia

24	148.4	6.9	150	2 AAV16207	Aav16207 Part of t
c 25	55.4	2.6	2000	7 ADA71938	Ada71938 Rice gene
26	51.6	2.4	2205	8 ADA28999	Ada28999 DNA encod
27	49	2.3	896	2 AAQ68903	Aaq68903 PspA prot
28	49	2.3	946	2 AAX33124	Aax33124 Streptoco
29	49	2.3	957	2 AAT61726	Aat61726 Streptoco
30	49	2.3	1860	7 ACA49648	Aca49648 Prokaryot
31	49	2.3	1990	7 AAT61725	Ast61725 Streptoco
32	49	2.3	1990	7 ABX95373	Abx95373 S. pneumo
33	49	2.3	2085	2 AAQ78131	Aaq78131 Pneumococ
34	49	2.3	2085	2 AAT08979	Aat08979 DNA encod
35	49	2.3	2085	2 AAT07178	Aat07178 Pneumococ
36	49	2.3	2085	2 AAV33264	Aav33264 Pneumococ
37	49	2.3	2085	2 AAV33264	Aav33264 Streptoco
38	49	2.3	2085	2 AAZ25063	Aaz25063 Streptoco
39	49	2.3	2085	2 AAV84069	Aav84069 DNA encod
40	49	2.3	2085	2 AAX02012	Aax02012 S. pneumo
41	47.4	2.2	956	7 ABX95374	Abx95374 S. pneumo
42	47.4	2.2	2085	2 AAQ28674	Aaq28674 PspA gene
43	47.2	2.2	110000	6 ABA90521_07	Continuation (8 of
c 44	47	2.2	9373	6 ABL33291	Ab133291 Human imm
c 45	47	2.2	9373	6 ABK31317	Abk31317 Signal tr

ALIGNMENTS

RESULT 1
AAD02063
ID AAD02063 standard; DNA; 2156 BP.
XX
AC AAD02063;
XX
DT 15-SEP-2003 (revised)
DT 26-MAR-2001 (first entry)
XX
DE Chlamydia pneumoniae 76 kDa full-length protein encoding DNA.
XX
KW 76 kDa protein; bactericidal; diagnosis; prevention; treatment;
KW pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; ds.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 101..2056
FT FT /*tag= a
FT FT /product= "Chlamydia pneumoniae 76KDa protein"

WO200066739-A2.

09-NOV-2000.

03-MAY-2000; 2000WO-CA000511.

03-MAY-1999; 99US-0132270P.

30-JUN-1999; 99US-0141276P.

(AVET) AVENTIS PASTEUR LTD.

Murdin AD, Oomen RP, Wang J, Dunn P;

WPI; 2000-687542/67.

P-PSDB; AAY71954.

Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful for vaccinating against Chlamydia infections.

Claim 2a; Fig 1; 112pp; English.

The present sequence is a DNA coding for Chlamydia pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the diagnosis,

Original

CC	prevention and treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory tract disease, bronchitis, sinusitis and acute respiratory disease such as cough, sore throat, hoarseness, fever; and abnormal chest sounds on auscultation). C. pneumoniae sequence is also used as vaccines for immunising humans against diseases caused by C. pneumoniae. (Updated on 15-SEP-2003 to standardise OS field)
XX	Sequence 2156 BP; 672 A; 461 C; 471 G; 552 T; 0 U; 0 Other;
SQ	
	Query Match 100.0%; Score 2156; DB 3; Length 2156; Best Local Similarity 100.0%; Pred. No. 0; Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATAAATCTTTAAACACAGCTCGCATTAATTAATTAAGTAGAGAGCTTTTTTTTATTTTTT 60
DB	1 ATAAATCTTTAAACACAGCTCGCATTAATTAATTAAGTAGAGAGCTTTTTTTTATTTTTT 60
QY	61 ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
DB	61 ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
QY	121 AGSTCTCTATAGAGAAACAGACGACACCTCCCGCAGATCTTCTGCTCAAGGATTTGA 180
DB	121 AGSTCTCTATAGAGAAACAGACGACACCTCCCGCAGATCTTCTGCTCAAGGATTTGA 180
QY	181 GGCGAGTCGAGCAAAATAAGAGTGGGAAGCTCAAGAGATAGAGGTGGGAAGCTAAGCC 240
DB	181 GGCGAGTCGAGCAAAATAAGAGTGGGAAGCTCAAGAGATAGAGGTGGGAAGCTAAGCC 240
QY	241 TAAAGAACTAAGACCGATTTCTGTAGAGGATGGAGCATCTTGGGTTCTGCAGTGAATGC 300
DB	241 TAAAGAACTAAGACCGATTTCTGTAGAGGATGGAGCATCTTGGGTTCTGCAGTGAATGC 300
QY	301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAG 360
DB	301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAG 360
QY	361 CAGATCTGCAGACGTGGATCTAACGACGACGCGACCTTACGCTCTCCACCCAGCTT 420
DB	361 CAGATCTGCAGACGTGGATCTAACGACGACGCGACCTTACGCTCTCCACCCAGCTT 420
QY	421 TGATGATTATAAGACTCAAGCGCAACACGTTTACGATATCTTTTACCTCAACATCACT 480
DB	421 TGATGATTATAAGACTCAAGCGCAACACGTTTACGATATCTTTTACCTCAACATCACT 480
QY	481 AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGCTCACTAATATAAGAGATAC 540
DB	481 AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGCTCACTAATATAAGAGATAC 540
QY	541 AGCGGCTACTGATAGGAAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGC 600
DB	541 AGCGGCTACTGATAGGAAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGC 600
QY	601 AGTTAAAGTTGGCGGCAAAATTAAGAAATAGCAATAGCAAAATAGCTTCGGATPAACAGAGCGAT 660
DB	601 AGTTAAAGTTGGCGGCAAAATTAAGAAATAGCAATAGCAAAATAGCTTCGGATPAACAGAGCGAT 660
QY	661 TCTTGACTCTTTAGGTAAACTGACTTCTCTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
DB	661 TCTTGACTCTTTAGGTAAACTGACTTCTCTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
QY	721 TGTAGCAAAACAATAACAGAGCTGAGCTTCTTAAAGAGATGCAAGATPAACCCAGTAGT 780
DB	721 TGTAGCAAAACAATAACAGAGCTGAGCTTCTTAAAGAGATGCAAGATPAACCCAGTAGT 780
QY	781 CCCAGGGAACCGCTGCAATTTCTCAATCTTTAGTGTAGATCAGAGATGCTTACAGCGAC 840
DB	781 CCCAGGGAACCGCTGCAATTTCTCAATCTTTAGTGTAGATCAGAGATGCTTACAGCGAC 840
QY	841 ACAGATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTGCAGGACGAAAGCGCTAG 900
DB	841 ACAGATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTGCAGGACGAAAGCGCTAG 900
QY	901 TGGAGCTCTAGAAAATGCTAAATCTAATAACAGTATATAAGCAACATAGATTTCAGCTAAAGC 960
DB	901 TGGAGCTCTAGAAAATGCTAAATCTAATAACAGTATATAAGCAACATAGATTTCAGCTAAAGC 960
QY	961 AGCAATCGCTACTGCTAAGACACAAAATAGCTGAAGCTCAGAAAAAGTTTCCCGACTCTCC 1020
DB	961 AGCAATCGCTACTGCTAAGACACAAAATAGCTGAAGCTCAGAAAAAGTTTCCCGACTCTCC 1020
QY	1021 AATTCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCAA 1080
DB	1021 AATTCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATCAA 1080
QY	1081 ACCTGCAGATGGTCTGATGTTCCAAATCCAGGAACACAGTTCGAGGCTCCAAAGCAACA 1140
DB	1081 ACCTGCAGATGGTCTGATGTTCCAAATCCAGGAACACAGTTCGAGGCTCCAAAGCAACA 1140
QY	1141 AGGAAGTAGTATTGGTAGTATTGGTATTTCCATGCTGTAGATGATGCTGAAATGAGAC 1200
DB	1141 AGGAAGTAGTATTGGTAGTATTGGTATTTCCATGCTGTAGATGATGCTGAAATGAGAC 1200
QY	1201 CGCTTCCATTTTGATGCTGGGTTTCGTCAGATGATTCAATGCTCAATACGGAATATCC 1260
DB	1201 CGCTTCCATTTTGATGCTGGGTTTCGTCAGATGATTCAATGCTCAATACGGAATATCC 1260
QY	1261 TGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGAGCAGCGAAAGCGCTGG 1320
DB	1261 TGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGAGCAGCGAAAGCGCTGG 1320
QY	1321 AGATGACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 1380
DB	1321 AGATGACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 1380
QY	1381 TAAAGCTGGGCAACAAACAGGCGATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGTGT 1440
DB	1381 TAAAGCTGGGCAACAAACAGGCGATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGTGT 1440
QY	1441 TGTAGCGCAGAGTTCTCCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAAGAACAGCT 1500
DB	1441 TGTAGCGCAGAGTTCTCCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAAGAACAGCT 1500
QY	1501 TTACAAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGA 1560
DB	1501 TTACAAGACCTCAAAATCTACAGGTTCTGATTATAAAACACAGATATCAGCAGGTTATGA 1560
QY	1561 TGCTTTACAAATCCATCAATGATGCTATGTTAGGCGCAGAAATGATGCGACTCGTGATGT 1620
DB	1561 TGCTTTACAAATCCATCAATGATGCTATGTTAGGCGCAGAAATGATGCGACTCGTGATGT 1620
QY	1621 GATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAAAGC 1680
DB	1621 GATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAAAGC 1680
QY	1681 TCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTTTCGCAATAGCAGAAC 1740
DB	1681 TCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTTTCGCAATAGCAGAAC 1740
QY	1741 TCCTTGGAGATCTTATAGTCAAGTTTCGGCACTCAATCTGTAATGCGAGATCACTCCAGTC 1800
DB	1741 TCCTTGGAGATCTTATAGTCAAGTTTCGGCACTCAATCTGTAATGCGAGATCACTCCAGTC 1800
QY	1801 GAATCCTCAAGCGAATTAATAGGAGATCAGACAAAAGCTTTACATCGCGACTGCAAAAGCC 1860
DB	1801 GAATCCTCAAGCGAATTAATAGGAGATCAGACAAAAGCTTTACATCGCGACTGCAAAAGCC 1860
QY	1861 TCCACAGTTTGGCTATCGTTATGTGCAACTTTCTAATGACTCTACACAGAGTTCTATAGC 1920
DB	1861 TCCACAGTTTGGCTATCGTTATGTGCAACTTTCTAATGACTCTACACAGAGTTCTATAGC 1920
QY	1921 TAAATTAGAAAAGTTTGTGCTGAAGGATCTAGSACAGAGCTGAAATAAAGCACTTTTC 1980
DB	1921 TAAATTAGAAAAGTTTGTGCTGAAGGATCTAGSACAGAGCTGAAATAAAGCACTTTTC 1980
QY	1981 CTTTGAACAGAACTCCCTGTTTATTTACGAGGTCGCTGCTCAATATCGGCTCTCTATATTC 2040

Db	1981	CTTTGAACGAACTCCTTGTATTCAGCAGTGTGCTCAATATCGGCTCTCTATATTC	2040
Qy	2041	TGGTTATCTCCAAATACACACCTAAGTGTTCGTTTGGAGAGATTATTATGCTTTGGT	2100
Db	2041	TGGTTATCTCCAAATACACACCTAAGTGTTCGTTTGGAGAGATTATTATGCTTTGGT	2100
Qy	2101	AAGGCTTTGTTGAGGCTTACCAACACACTAGACGATCTTCAATAATAAAGA	2156
Db	2101	AAGGCTTTGTTGAGGCTTACCAACACACTAGACGATCTTCAATAATAAAGA	2156
RESULT 2			
AAx91990_08			
Continuation (9 of 13) of AAx91990 from base 800001 (Nucleotide sequence of the complete			
WP Sequence split into 13 fragments LOCUS AAx91990 Accession AAx91990			
WP	Fragment Name	Begin	End
WP	AAx91990_00	1	110000
WP	AAx91990_01	100001	210000
WP	AAx91990_02	200001	310000
WP	AAx91990_03	300001	410000
WP	AAx91990_04	400001	510000
WP	AAx91990_05	500001	610000
WP	AAx91990_06	600001	710000
WP	AAx91990_07	700001	810000
WP	AAx91990_08	800001	910000
WP	AAx91990_09	900001	1010000
WP	AAx91990_10	1000001	1110000
WP	AAx91990_11	1100001	1210000
WP	AAx91990_12	1200001	1230025
Query Match			
Best Local Similarity 99.4%; Score 2143.4; DB 2; Length 110000;			
Matches 2155; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	1	ATAAAATCTTTAAACAGGCTCGCATTAATTTAGTGTGAGAGCTTTTATTTT	60
Db	28698	ATAAAATCTTTAAACAGGCTCGCATTAATTTAGTGTGAGAGCTTTTATTTT	28639
Qy	61	ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	120
Db	28638	ATAATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	28579
Qy	121	AGGTCCTATAGACGAAACAGAACGACACCTCCGCGAGATCTTCTGCTCAAGGATTGGA	180
Db	28578	AGGTCCTATAGACGAAACAGAACGACACCTCCGCGAGATCTTCTGCTCAAGGATTGGA	28519
Qy	181	GGCAGTGCAGCAATATAGAGTGGGAAGCTCAAGAAATAGAGTGGGAAGCTAAGCC	240
Db	28518	GGCAGTGCAGCAATATAGAGTGGGAAGCTCAAGAAATAGAGTGGGAAGCTAAGCC	28459
Qy	241	TAAAGAACTAAGACCGATTCGTAGACGATGGAGCATCTTGCGTCTCGAGTGAATGC	300
Db	28458	TAAAGAACTAAGACCGATTCGTAGACGATGGAGCATCTTGCGTCTCGAGTGAATGC	28399
Qy	301	TCTCATAGTCTGCAGATAAGCTGGGTATTTCTTAGTAACAGCTCGTCTTCTACTAG	360
Db	28398	TCTCATAGTCTGCAGATAAGCTGGGTATTTCTTAGTAACAGCTCGTCTTCTACTAG	28339
Qy	361	CAGATCTGCAGACGTGAGCTCAACAGACGACGACCTACGCTCTCCACCCAGTT	420
Db	28338	CAGATCTGCAGACGTGAGCTCAACAGACGACGACCTACGCTCTCCACCCAGTT	28279
Qy	421	TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT	480
Db	28278	TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACT	28219
Qy	481	AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGCTACTAATAATAAGGATAC	540
Db	28218	AGCTGACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTGCTACTAATAATAAGGATAC	28159
Qy	541	AGCGGCTACTGATAGGAAACCGCAATGCTGCGGAGTGGGAATCAAGAAATGCCGATGC	600

Db	28158	AGCGGCTACTGATAGGAAACCGCAATCGCTGCGAGTGGGAAACTAAGATGCCGATGC	28099
Qy	601	AGTTAAAGTTGCGCGCAAAATTACAGAAATTACGAAATATATCTTCGGATAACCAAGCGAT	660
Db	28098	AGTTAAAGTTGCGCGCAAAATTACAGAAATTAGGAAATATATCTTCGGATAACCAAGCGAT	28039
Qy	661	TCATTGACCTTTAGGTAATGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC	720
Db	28038	TCATTGACCTTTAGGTAATGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATC	27979
Qy	721	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	780
Db	27978	TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	27919
Qy	781	CCCAGGAAACCGCTGCAATGCTCAATCTTAACTTTAGTTGATCAGACAGATGTACAGCGAC	840
Db	27918	CCCAGGAAACCGCTGCAATGCTCAATCTTAACTTTAGTTGATCAGACAGATGTACAGCGAC	27859
Qy	841	ACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTCAGGACAGACGCTAG	900
Db	27858	ACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTCAGGACAGACGCTAG	27799
Qy	901	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGCTAAAGC	960
Db	27798	TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTGAGCTAAAGC	27739
Qy	961	AGCAATCGCTACTGTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC	1020
Db	27738	AGCAATCGCTACTGTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC	27679
Qy	1021	AATTCTTTCAAGAGCGGAACAAATGTTAAATACAGGCTGAGAAAGATCTTAAATAATCAA	1080
Db	27678	AATTCTTTCAAGAGCGGAACAAATGTTAAATACAGGCTGAGAAAGATCTTAAATAATCAA	27619
Qy	1081	ACCTGACAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACA	1140
Db	27618	ACCTGACAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACA	27559
Qy	1141	AGGAAGTAGTATTGGTATGATTCGTTGTTCCATGCTGTAGATGATGCTGAAAATGAGAC	1200
Db	27558	AGGAAGTAGTATTGGTATGATTCGTTGTTCCATGCTGTAGATGATGCTGAAAATGAGAC	27499
Qy	1201	CGCTTCATTTTGATGTCGTTGTTCCATGATGATTCATGATGATTCATGCGGAAATCC	1260
Db	27498	CGCTTCATTTTGATGTCGTTGTTCCATGATGATTCATGATGATTCATGCGGAAATCC	27439
Qy	1261	TGATTCCTAAGCTGCCCAACAGGAGCTCCGAGCAACAGCTAGACGACGAAAGCCGCTGG	1320
Db	27438	TGATTCCTAAGCTGCCCAACAGGAGCTCCGAGCAACAGCTAGACGACGAAAGCCGCTGG	27379
Qy	1321	AGATGACAGTCTGCTGCGAGGCTGCGAGATGCTCAGAAAAGCTTTAGAACGGCTTAGG	1380
Db	27378	AGATGACAGTCTGCTGCGAGGCTGCGAGATGCTCAGAAAAGCTTTAGAACGGCTTAGG	27319
Qy	1381	TAAAGCTGGGCAACAAACAGGGCTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTG	1439
Db	27318	TAAAGCTGGGCAACAAACAGGGCTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTG	27259
Qy	1440	TTGTGAGCGCAGAGTTCTCCCGCTGCACAAAGTTCTATAGGGTCACTGTAAAAACAGC	1499
Db	27258	TTGTGAGCGCAGAGTTCTCCCGCTGCACAAAGTTCTATAGGGTCACTGTAAAAACAGC	27199
Qy	1500	TTTACAGAGCTCAAAATCTACAGGTTCTGATTTATAAACACAGATATCAGCAGGTTATG	1559
Db	27198	TTTACAGAGCTCAAAATCTACAGGTTCTGATTTATAAACACAGATATCAGCAGGTTATG	27139
Qy	1560	ATGCTTACAATCCATCAATGATGCTATGTTAGGCGACGAAATGATGCGACTCTGATG	1619
Db	27138	ATGCTTACAATCCATCAATGATGCTATGTTAGGCGACGAAATGATGCGACTCTGATG	27079
Qy	1620	TGATTAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAAG	1679
Db	27078	TGATTAACAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAAG	27019

QY 1680 CTCGAGCACCAGAAAAAAGAGATCAAGCCCTCGCTAGGTGGTCTTCTGCAATAGCAGAA 1739
 Db 27018 CTCGAGCACCAGAAAAAAGAGATCAAGCCCTCGCTAGGTGGTCTTCTGCAATAGCAGAA 26959
 QY 1740 CTCCTCGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCGAGATCATCCAGT 1799
 Db 26958 CTCCTCGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCGAGATCATCCAGT 26899
 QY 1800 CGAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGAGTGAACAAAGC 1859
 Db 26898 CGAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGAGTGAACAAAGC 26839
 QY 1860 CTCACAGTTTGGCTATCCTTATGTGCAACTTCTTAATGACTCTACACAGAAGTTCATAG 1919
 Db 26838 CTCACAGTTTGGCTATCCTTATGTGCAACTTCTTAATGACTCTACACAGAAGTTCATAG 26779
 QY 1920 CTAAATGAAAGTTTGGTGTGGAAGGATCTAGGACAGCTGAATATAAAGCACTTT 1979
 Db 26778 CTAAATGAAAGTTTGGTGTGGAAGGATCTAGGACAGCTGAATATAAAGCACTTT 26719
 QY 1980 CTTTGAACGAACTCTCTTGTATTTACAGAGTCTGCTGATCAATATCGGCTCTCTATAT 2039
 Db 26718 CTTTGAACGAACTCTCTTGTATTTACAGAGTCTGCTGATCAATATCGGCTCTCTATAT 26659
 QY 2040 CTGGTTATCTCCAATAACAACACCTAAGTGTGTTGGAGAGATTAATATGCTGTTGG 2099
 Db 26658 CTGGTTATCTCCAATAACAACACCTAAGTGTGTTGGAGAGATTAATATGCTGTTGG 26599
 QY 2100 TAAAGCCTTTGTTAGGCTTACCAACACCTAGAACGATCTTCAATAATAAAGA 2156
 Db 26598 TAAAGCCTTTGTTAGGCTTACCAACACCTAGAACGATCTTCAATAATAAAGA 26542

RESULT 3

AAS57031

ID AAS57031 standard; DNA; 1956 BP.

XX AAS57031;

XX 11-SEP-2003 (revised)

XX 16-JAN-2002 (first entry)

XX C. pneumoniae DNA encoding the CT622 homologue CPn0728.

XX Chlamydia; ds; sexually transmitted disease; PID; antibacterial;

XX pelvic inflammatory disease; antigen; trachoma; gynecological;

XX acute respiratory tract infection; atherosclerosis; male infertility;

XX coronary heart disease.

XX Chlamydophila pneumoniae.

XX NC024018.1379-A2.

XX 01-NOV-2001.

XX 23-APR-2001; 2001WO-US013081.

XX 21-APR-2000; 2000US-0198853P.

XX 20-JUL-2000; 2000US-0219752P.

XX (CORI-) CORIXA CORP.

XX Bhatia A, Probst P, Stromberg EJ;

XX WPI; 2001-616771/71.

XX P-PSDB; AAU38899.

XX New polynucleotide for treating Chlamydia infections encodes a

XX polynucleotides containing an immunogenic portion of a Chlamydia antigen.

XX Disclosure; Page 143; 208pp; English.

XX

The invention relates to isolated polynucleotide encoding at least a partial Chlamydia protein which is an antigenic fragment, or the complements, fragments, homologues and variants, and antibodies raised against the antigenic proteins (or fragments). The nucleic acids, proteins and antibodies are used to diagnose and treat Chlamydia infections (e.g. a sexually transmitted disease, pelvic inflammatory disease (PID), acute respiratory tract infection, trachoma, atherosclerosis and coronary heart disease) in a patient, and in the treatment of male infertility. The compounds of the invention are also useful for detecting the presence of Chlamydia in a patient, and stimulating and/or expanding T cells specific for a Chlamydia protein. The present sequence encodes a Chlamydia antigen. (Updated on 11-SEP-2003 to standardise OS field)

Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

Query Match 90.7%; Score 1956; DB 4; Length 1956;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 160

Db 1 ATGGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 60

QY 161 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATAGAGTGGGAAGCTCAAGAATA 220

Db 61 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATAGAGTGGGAAGCTCAAGAATA 120

QY 221 GCAGTGGCGAAGCTAAGCCCTAAAGAAATCTAAGACCGATTTCTAGAGCGATGGAGCATC 280

Db 121 GCAGTGGCGAAGCTAAGCCCTAAAGAAATCTAAGACCGATTTCTAGAGCGATGGAGCATC 180

QY 281 TTGGTCTGTCAGTGAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATTTGCTTCTAGT 340

Db 181 TTGGTCTGTCAGTGAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATTTGCTTCTAGT 240

QY 341 AACAGCTGCTCTTCTACTAGCAGATCTGCAGCTGGACTCAACGACGACGACCGCAGCT 400

Db 241 AACAGCTGCTCTTCTACTAGCAGATCTGCAGCTGGACTCAACGACGACGACCGCAGCT 300

QY 401 ACGCTCTCTCCACCCAGCTTTGATTAAGACTCAAGCGCAAAACAGCTTACGATACT 460

Db 301 ACGCTCTCTCCACCCAGCTTTGATTAAGACTCAAGCGCAAAACAGCTTACGATACT 360

QY 461 ATCTTTACCTCAACATCACTAGCTGACATACAGCTGCTTTGGTGGAGCTCCAGATGCT 520

Db 361 ATCTTTACCTCAACATCACTAGCTGACATACAGCTGCTTTGGTGGAGCTCCAGATGCT 420

QY 521 GTCCTAATAATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGG 580

Db 421 GTCCTAATAATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGG 480

QY 581 GAAACTAAGATGCGGATGAGTTAAAGTTGGCGCGCAAAATACAGAATTAAGCGAAATAT 640

Db 481 GAAACTAAGATGCGGATGAGTTAAAGTTGGCGCGCAAAATACAGAATTAAGCGAAATAT 540

QY 641 GCTTCGGATTAACCAAGCGATTTCTGACTCTTTAGTAACTGACTTCTCTCGACTCTTA 700

Db 541 GCTTCGGATTAACCAAGCGATTTCTGACTCTTTAGTAACTGACTTCTCTCGACTCTTA 600

QY 701 CAGGCTGCTCTTCTTCCAAATCTGTAGCAAAACAAATAAAGCAGCTGAGCTTCTTAAAGAG 760

Db 601 CAGGCTGCTCTTCTTCCAAATCTGTAGCAAAACAAATAAAGCAGCTGAGCTTCTTAAAGAG 660

QY 761 ATGCAAGATTAACCCAGTAGTCCCGAGGAAACCGCTCGCAATTCCTAATCTTTAGTTGAT 820

Db 661 ATGCAAGATTAACCCAGTAGTCCCGAGGAAACCGCTCGCAATTCCTAATCTTTAGTTGAT 720

QY 821 CAGACATGCTCAGCGACACAGATAGAGAAATGGAATCGGATTAGGGATGCATAT 880

Db 721 CAGACATGCTCAGCGACACAGATAGAGAAATGGAATCGGATTAGGGATGCATAT 780

QY 881 TTTGCGAGCAGAACGCTAGTGGAGCTGTAGAAAAATGCTAAATCTAATAACAGTATAAGC 940

Db 781 TTTGAGGACGAGACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 840
Qy 941 AACATAGATTACGCTAAAGCAGCAATCGCTAGCTAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTACGCTAAAGCAGCAATCGCTAGCTAGACACAAATAGCTGAGCTCAG 900
Qy 1001 AAAAGTTCCCGACTCTCCAAATCTTCAAGAACGGGAACAAATGGTAAATACAGGCTCAG 1060
Db 901 AAAAGTTCCCGACTCTCCAAATCTTCAAGAACGGGAACAAATGGTAAATACAGGCTCAG 960
Qy 1061 AAGATCTTAAATAATCAAACTCGCAGATGGTCTGATGTTCCAAATCCAGGAACATA 1120
Db 961 AAGATCTTAAATAATCAAACTCGCAGATGGTCTGATGTTCCAAATCCAGGAACATA 1020
Qy 1121 GTTGAGGCTCCAGCAACAGGAGTAGTATGGTAGTATTCGGTTCCTCATGCTGTTA 1180
Db 1021 GTTGAGGCTCCAGCAACAGGAGTAGTATGGTAGTATTCGGTTCCTCATGCTGTTA 1080
Qy 1181 GATGATGCTGAAATGAGACCGCTTCCATTTTGATGCTGGGTTTCTGTCAGATGATTCAC 1240
Db 1081 GATGATGCTGAAATGAGACCGCTTCCATTTTGATGCTGGGTTTCTGTCAGATGATTCAC 1140
Qy 1241 ATGTTCAATACGGAATTCCTGATCTCAAGCTGCCCAACAGGAGCTCGAGCAAGCT 1300
Db 1141 ATGTTCAATACGGAATTCCTGATCTCAAGCTGCCCAACAGGAGCTCGAGCAAGCT 1200
Qy 1301 AGAGCAGGAAGCGCTGGAGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAA 1360
Db 1201 AGAGCAGGAAGCGCTGGAGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAA 1260
Qy 1361 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGGTAAGCTGGGCAACAAACAGGCGATCAATGCTTTAGGA 1320
Qy 1421 CAGATCGTCTGCTGTTGTCGCGCAGAGTCTCTCCGCTGAGCAAGTCTTATA 1480
Db 1321 CAGATCGTCTGCTGTTGTCGCGCAGAGTCTCTCCGCTGAGCAAGTCTTATA 1380
Qy 1481 GGGTCATCTGTAACACAGCTTTACAGACCTCAAAATCTACAGTCTCTGATTATAAACA 1540
Db 1381 GGGTCATCTGTAACACAGCTTTACAGACCTCAAAATCTACAGTCTCTGATTATAAACA 1440
Qy 1541 CAGATACAGAGGTTATGATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1600
Db 1441 CAGATACAGAGGTTATGATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1500
Qy 1601 AATGATCGACTCTGATGATGATAAACAATGATGATGATGATGATGATGATGATGATG 1660
Db 1501 AATGATCGACTCTGATGATGATAAACAATGATGATGATGATGATGATGATGATGATG 1560
Qy 1661 CCTAGACAGCAACAGAGCTTCGAGGACCCAGAAAAAAGATCAAGCCCTCGCTAGGGTG 1720
Db 1561 CCTAGACAGCAACAGAGCTTCGAGGACCCAGAAAAAAGATCAAGCCCTCGCTAGGGTG 1620
Qy 1721 ATTCTGCAATAGCAGAACTCTTGGAGATGCTTATGATCAAGTTTCGCACTACAAATCT 1780
Db 1621 ATTCTGCAATAGCAGAACTCTTGGAGATGCTTATGATCAAGTTTCGCACTACAAATCT 1680
Qy 1781 GTAATGAGATCATCCAGTGAATCTTCAAGCGAATATAGGAGATCAGCAAAAGCTT 1840
Db 1581 GTAATGAGATCATCCAGTGAATCTTCAAGCGAATATAGGAGATCAGCAAAAGCTT 1740
Qy 1841 ACATCGGAGTACAAAGCCCTCCAGTTCGGCTATCCTATGTCGAATCTTCTAATGAC 1900
Db 1741 ACATCGGAGTACAAAGCCCTCCAGTTCGGCTATCCTATGTCGAATCTTCTAATGAC 1800
Qy 1901 TCTACAGAAAGTTCATAGCTTAATTAAGATTTGTTGCTGAAGATCTAGACAGCA 1960
Db 1801 TCTACAGAAAGTTCATAGCTTAATTAAGATTTGTTGCTGAAGATCTAGACAGCA 1860
Qy 1961 GCTGAATAAAGCACCTTCTCTTGAACGAACTCCCTGTTTATTTACAGAGGCTGCTGTC 2020

Db 1861 GCTGAATAAAGCACCTTCTCTTGAAGCAACTCTTGTATTATTACAGAGGTCGTGTC 1920
Qy 2021 AATATCGGCTCTCTATATATTCGTGTTATCTCAATAA 2056
Db 1921 AATATCGGCTCTCTATATATTCGTGTTATCTCAATAA 1956

RESULT 4

ABL92612

ID ABL92612 standard; DNA; 1956 BP.

XX ABL92612;

AC

XX 29-AUG-2003 (revised)

DT 05-JUN-2002 (first entry)

XX

XX Chlamydia pneumoniae DNA sequence SEQ ID NO:385.

DE

XX Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;

KW antibacterial; immunostimulant; immune response;

KW Chlamydia-specific T-cell response; gene; ds.

XX

OS Chlamydia pneumoniae.

XX

XX WD2020208267-A2.

XX

XX 31-JAN-2002.

XX

XX 20-JUL-2001; 2001WO-US023121.

XX

XX 20-JUL-2000; 2000US-00620412.

XX

XX 23-APR-2001; 2001US-00841132.

XX

XX (CORI) CORIXA CORP.

XX

XX Fling SP, Skeiky YAM, Probst P, Bhatia A;

XX

XX WPI; 2002-179901/23.

XX

XX Novel compositions comprising Chlamydia CapI protein and its use in the treatment of Chlamydia infection.

XX

XX Disclosure; Page 336-337; 537pp; English.

XX

The present invention describes compositions comprising a Chlamydia CapI protein and methods for the diagnosis and therapy of Chlamydia infection. Chlamydia DNA and protein sequences from the present invention can have antibacterial and immunostimulant activities, and can be used in vaccines. Compounds from the present invention can be used for eliciting an immune response, specifically stimulating a Chlamydia-specific T-cell response or inhibiting the development of a Chlamydia infection in an animal. Methods from the present invention can be used: for detecting the presence of Chlamydia in a patient; to stimulate and/or expand T cells specific for a Chlamydia protein; and for treatment of a Chlamydia infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent sequences used in the exemplification of the present invention. (Updated on 29-AUG-2003 to standardise OS field)

XX

SQ Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

XX

XX Query Match 90.7%; Score 1956; DB 6; Length 1956;

XX

XX Best Local Similarity 100.0%; Pred. No. 0;

XX

XX Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

101 ATGGTTAATCTTATTTGGTCCAGGTCTTATAGAGAAACAGACGACCTCCCGAGAT 160

Db

1 ATGGTTAATCTTATTTGGTCCAGGTCTTATAGAGAAACAGACGACCTCCCGAGAT 60

Qy

161 CTCTTCTGCTCAAGATTGGAGGCGAGTGCAGCAAAATAAGAGTCGGAAGCTCAAGAATA 220

Db

61 CTCTTCTGCTCAAGATTGGAGGCGAGTGCAGCAAAATAAGAGTCGGAAGCTCAAGAATA 120

Qy

221 GCAGGTGCGGAAGCTAAGCCTAAAGAACTCTAGACCGGATTCTGTAGAGCATGGAGCATC 280

Db 121 GCAGTGGGAAAGCTAAGCTAAAGATCTAAGACCGGATCTGTAGACCGATGGAGCATC 180
QY 281 TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGT 240
QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACAGACGACACCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACAGACGACACCGACCT 300
QY 401 AGCGTCTCTCAACCCAGCTTTGATGATTAAGATCTAAGCGCAACAGCTTACGATCT 460
Db 301 AGCGTCTCTCAACCCAGCTTTGATGATTAAGATCTAAGCGCAACAGCTTACGATCT 360
QY 461 ATCTTTACTCTCAACATCACTAGCTCACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
Db 361 ATCTTTACTCTCAACATCACTAGCTCACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
QY 521 GTCACATAATAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGGAGTGG 580
Db 421 GTCACATAATAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGGAGTGG 480
QY 581 GAACTAGATAGCCGATGCGAGTTAAAGTTGGCGCGCAAAATACAGATTAAGCAATAT 640
Db 481 GAACTAGATAGCCGATGCGAGTTAAAGTTGGCGCGCAAAATACAGATTAAGCAATAT 540
QY 641 GCTTCGGATAACCAAGCGATCTTGACTCTTTAGTAAATGACTTCTCTCGACCTCTTA 700
Db 541 GCTTCGGATAACCAAGCGATCTTGACTCTTTAGTAAATGACTTCTCTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCTCAATCTGTAGCAACAAATAACAAAGAGCTGAGCTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCAATCTGTAGCAACAAATAACAAAGAGCTGAGCTCTTAAAGAG 660
QY 761 ATGCAAGATAACCCAGTGTCCCGGGAAGCGCTGCAATGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTGTCCCGGGAAGCGCTGCAATGCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCATAT 880
Db 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCATAT 780
QY 881 TTTCAGGACGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 940
Db 781 TTTCAGGACGAAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGC 840
QY 941 AACATAGATTACGCTAAAGCAGCAATCGCTACTGTCTAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTACGCTAAAGCAGCAATCGCTACTGTCTAGACACAAATAGCTGAGCTCAG 900
QY 1001 AAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATCAGGCTGAG 1060
Db 901 AAAAGTTCCCGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAAATCAGGCTGAG 960
QY 1061 AAAAGTTCTTAAATAATCAACCTCGAGTGTCTGATGTTTCAATCCAGGAACTACA 1120
Db 961 AAAAGTTCTTAAATAATCAACCTCGAGTGTCTGATGTTTCAATCCAGGAACTACA 1020
QY 1121 GTTGGAGGCTCCAGCAACAGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGGCTCCAGCAACAGGAAGTAGTATTGGTAGTATTTCGTTTCCATGCTGTTA 1080
QY 1181 GATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGCTGGTTCGTGAGATGATTCAC 1240
Db 1081 GATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGCTGGTTCGTGAGATGATTCAC 1140
QY 1241 ATGTTCAATACGGAATTCCTGATTTCTAAGCTGCCCAACAGGAGCTCGACGACAACT 1300
Db 1141 ATGTTCAATACGGAATTCCTGATTTCTAAGCTGCCCAACAGGAGCTCGACGACAACT 1200
QY 1301 AGACGACGAAAGCCGCTGGAGATGACAGTGTCTGTGAGCGCTGCGAGATGCTCAGAAA 1360

Db 1201 AGACGCGAAAAGCCGCTGGAGATGACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAA 1260
QY 1361 GCTTTAGAGCGGCTCTAGCTAAAGCTGGGCAACAACAGGCGCATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGGCTCTAGCTAAAGCTGGGCAACAACAGGCGCATCTCAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGTGTTGTGAGCGCAGGAGTTCTCCCGCTGAGCAAGTTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGTGTTGTGAGCGCAGGAGTTCTCCCGCTGAGCAAGTTCTATA 1380
QY 1481 GGGTCATCTGTAAACAAGCTTTTCAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1540
Db 1381 GGGTCATCTGTAAACAAGCTTTTCAAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1440
QY 1541 CAGATATCAGCAGGTTATGATGCTTCAAAATCCATCAATGATGCTATGTTAGGCGCAGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGCTTCAAAATCCATCAATGATGCTATGTTAGGCGCAGA 1500
QY 1601 AATGATCGACTCTGTATGTTGATAAACAATGTAAAGTACCCCGCTCTCACACGATCCGTT 1660
Db 1501 AATGATCGACTCTGTATGTTGATAAACAATGTAAAGTACCCCGCTCTCACACGATCCGTT 1560
QY 1661 CCTAGACGACGAAACAGAGCTCGAGGACCGAAGAAACAGATCAAGCCCTCGTAGGGTG 1720
Db 1561 CCTAGACGACGAAACAGAGCTCGAGGACCGAAGAAACAGATCAAGCCCTCGTAGGGTG 1620
QY 1721 ATTTCTGGCAATAGCAGAACTCTTTGGAGATGCTCTATAGTCAAGTTTCGGCACTACAATCT 1780
Db 1621 ATTTCTGGCAATAGCAGAACTCTTTGGAGATGCTCTATAGTCAAGTTTCGGCACTACAATCT 1680
QY 1781 GTAAATGAGATCATCCAGTCAAGTCCCTCAAGCGAATTAATGAGGAGATCAGCAAAAGCTT 1840
Db 1681 GTAAATGAGATCATCCAGTCAAGTCCCTCAAGCGAATTAATGAGGAGATCAGCAAAAGCTT 1740
QY 1841 ACATCGGACGTGACAAAGCCCTCCACAGTCTTGGCTATCTTATGTCGAATCTTAAATCAC 1900
Db 1741 ACATCGGACGTGACAAAGCCCTCCACAGTCTTGGCTATCTTATGTCGAATCTTAAATCAC 1800
QY 1901 TCTACACAGAAATTCATAGCTTAAATAGAAAGTTTGTTCCTGAAGGATCTAGGACAGCA 1960
Db 1801 TCTACACAGAAATTCATAGCTTAAATAGAAAGTTTGTTCCTGAAGGATCTAGGACAGCA 1860
QY 1961 GCTGAAATAAAGACATCTTCTTTGAAACCGAATCTCTTGTGTTTATTCAGCAGGCTGCTGTC 2020
Db 1861 GCTGAAATAAAGACATCTTCTTTGAAACCGAATCTCTTGTGTTTATTCAGCAGGCTGCTGTC 1920
QY 2021 AATATCGGCTCTCTATATCTTGTGTTATCTCCAAATAA 2056
Db 1921 AATATCGGCTCTCTATATCTTGTGTTATCTCCAAATAA 1956

RESULT 5

ABL91190

ID ABL91190 standard; DNA; 1956 BP.

XX ABL91190;

XX AC ABL91190;

XX 29-AUG-2003 (revised)

DT 29-JUL-2002 (first entry)

XX Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.

XX Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.

KW human respiratory disease; antigen; immunogen; vaccine; diagnosis;

KW coronary artery disease; atherosclerosis; carotid artery stenosis; myocardial infarction;

KW cerebrovascular disease; aortic aneurysm; claudication; stroke;

KW strain CWF029; open reading frame; ORF; gene; ds.

XX Chlamydia pneumoniae.

OS Chlamydia pneumoniae.

XX Key Location/Qualifiers

FH 1..1956

FT CDS /*tag= a

FT XX
PN XX
XX XX
PD XX
PF XX
PR XX
PR 11-JUL-2000; 2000GB-00016363.
PR 21-JUL-2000; 2000GB-00017047.
PR 07-AUG-2000; 2000GB-00017983.
PR 18-AUG-2000; 2000GB-00020440.
PR 14-SEP-2000; 2000GB-00022583.
PR 10-NOV-2000; 2000GB-00027549.
PR 22-DEC-2000; 2000GB-00031706.
XX (CHIR-) CHIRON SPA.
XX Ratti G, Grandi G;
XX WPI; 2002-154726/20.
XX N-PSDB; ABB90532.
XX Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX Claim 5; Page 47-48; 364pp; English.
XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in the detection of
CC Chlamydia pneumoniae gene expression. The present sequence represents a
CC specifically claimed DNA which encodes a Chlamydia pneumoniae protein of
CC the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX SQ
Sequence 1956 BP; 506 A; 436 C; 441 G; 473 T; 0 U; 0 Other;
Query Match 90.7%; Score 1956; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 ATGGTTAATCCTATGCTCCAGTCCCTATAGACGAAACAGACACACCTCCCGGAGAT 160
Db 1 ATGGTTAATCCTATGCTCCAGTCCCTATAGACGAAACAGACACACCTCCCGGAGAT 60
QY 161 CTTTCTCTCAAGGATTGGAGGCGAGTGCAGCAATTAAGAGTGGGAGCTCAAGAGATA 220
Db 61 CTTTCTCTCAAGGATTGGAGGCGAGTGCAGCAATTAAGAGTGGGAGCTCAAGAGATA 120
QY 221 GCAGGTGGGAGTAAGCCCTAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATC 280
Db 121 GCAGGTGGGAGTAAGCCCTAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATC 180
QY 281 TTGCGTTCTCAGTGAATGCTCATGCTGTGAGTGAATGCTGATGCTTCTTCTAGT 340
Db 181 TTGCGTTCTCAGTGAATGCTCATGCTGTGAGTGAATGCTGATGCTTCTTCTAGT 240
QY 341 AACAGTCGCTTCTTCTACTAGCAGATTCGAGCGTGCAGCTCAACGACGAGCGGACCT 400
Db 241 AACAGTCGCTTCTTCTACTAGCAGATTCGAGCGTGCAGCTCAACGACGAGCGGACCT 300

QY 401 ACGCTCCTCCACCACCGTTTGATTAAGACTCAACGCGAAACAGCTTACGATACT 460
Db 301 ACGCTCCTCCACCACCGTTTGATTAAGACTCAACGCGAAACAGCTTACGATACT 360
QY 461 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCT 520
Db 361 ATCTTTACCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCT 420
QY 521 GTCACTAATATAAGGATACAGCGCTTACTGATGAGGAAACCGCAATCGCTGCGGAGTGG 580
Db 421 GTCACTAATATAAGGATACAGCGCTTACTGATGAGGAAACCGCAATCGCTGCGGAGTGG 480
QY 581 GAAACTAAGATGCGATGAGTGAAGTTGAGGCGCAAAATTAAGAAATAGCAAAATAT 640
Db 481 GAAACTAAGATGCGATGAGTGAAGTTGAGGCGCAAAATTAAGAAATAGCAAAATAT 540
QY 641 GCTTCGGATAACCAAGCGATTCTTGACTCTTTAGGTAAGCTGACTCTCTCGACTCTTA 700
Db 541 GCTTCGGATAACCAAGCGATTCTTGACTCTTTAGGTAAGCTGACTCTCTCGACTCTTA 600
QY 701 CAGCTGCTCTCTTCCCAATCTGTAGCAAAACAAATTAAGAGCTGAGCTTCTTAAGAG 760
Db 601 CAGCTGCTCTCTTCCCAATCTGTAGCAAAACAAATTAAGAGCTGAGCTTCTTAAGAG 660
QY 761 ATGCAAGATAAACCAGTAGTCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAAACCAGTAGTCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGATGCATAT 880
Db 721 CAGACAGATCTACAGCGACACAGATAGAGAAAGATGGAATCGATTAGGATGCATAT 780
QY 881 TTTTGCAGGACAGACGCTAGTGAAGCTGTAGAAAAGCTAAATCTAATAACATGATTAAGC 940
Db 781 TTTTGCAGGACAGACGCTAGTGAAGCTGTAGAAAAGCTAAATCTAATAACATGATTAAGC 840
QY 941 AACATAGATTCAGTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTCAGTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
QY 1001 AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGTTAATACAGGCTGAG 1060
Db 901 AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGTTAATACAGGCTGAG 960
QY 1061 AAAGATCTTAAATAATCAACCTGAGAGTGTCTGATGCTTCCAAATCCAGACACTACA 1120
Db 961 AAAGATCTTAAATAATCAACCTGAGATGTTCTGATGTTCCAAATCCAGACACTACA 1020
QY 1121 GTTGGAGGCTCCAAAGCAACAAAGGAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTA 1180
Db 1021 GTTGGAGGCTCCAAAGCAACAAAGGAGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTA 1080
QY 1181 GATGATGCTGAAATAGACCGCTTCCAAATTTGATGCTGGGTTTCTGATGATGATTCAC 1240
Db 1081 GATGATGCTGAAATAGACCGCTTCCAAATTTGATGCTGGGTTTCTGATGATGATTCAC 1140
QY 1241 ATGTTCAATACGAAATCTCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGACGACAGCT 1300
Db 1141 ATGTTCAATACGAAATCTCTGATTTCTCAAGCTGCCCAACAGGAGCTCCGACGACAGCT 1200
QY 1301 AGACGAGCAAAACCGCTCGAGATGACAGTGTGCTGTCAGCGCTGCAGATGCTCAGAAA 1360
Db 1201 AGACGAGCAAAACCGCTCGAGATGACAGTGTGCTGTCAGCGCTGCAGATGCTCAGAAA 1260
QY 1361 GCTTTAGAGCGGCTCTAGTAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGGCTCTAGTAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1480
Db 1321 CAGATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380

QY 1481 GGGTCATCTGTAAACACAGCTTTTACAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1540
DB 1391 GGGTCATCTGTAAACACAGCTTTTACAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1440
QY 1541 CAGTATCAGCAGGTTATGATGCTTTACAAATCCATCAATGATGCTATGTTAGGACGCA 1600
DB 1441 CAGTATCAGCAGGTTATGATGCTTTACAAATCCATCAATGATGCTATGTTAGGACGCA 1500
QY 1601 AATGATCGGATCGTGATGCTATGAACCAATCTAGTACCCCGCTCTCAGACGATCCGTT 1660
DB 1501 AATGATCGGATCGTGATGCTATGAACCAATCTAGTACCCCGCTCTCAGACGATCCGTT 1560
QY 1661 CCTAGACGACGACAGAGCTCGAGGACGAGAAACAGATCAAGCCCTCGCTAGGCTG 1720
DB 1561 CCTAGACGACGACAGAGCTCGAGGACGAGAAACAGATCAAGCCCTCGCTAGGCTG 1620
QY 1721 ATTTCTGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTGGCACTCAATCT 1780
DB 1621 ATTTCTGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTGGCACTCAATCT 1680
QY 1781 GTAATGCAATCATCCAGTCCGAATCTCAAGCGAATATGAGGATCAGACAAAGCTT 1840
DB 1681 GTAATGCAATCATCCAGTCCGAATCTCAAGCGAATATGAGGATCAGACAAAGCTT 1740
QY 1841 ACATCGGAGTGACAAAGCTCCAGGTTTGGCTATCTTATGTCGAATTTCTAATGAC 1900
DB 1741 ACATCGGAGTGACAAAGCTCCAGGTTTGGCTATCTTATGTCGAATTTCTAATGAC 1800
QY 1901 TCTACAGAGATTCATAGCTAAATAGAAAGTTTGGTGAAGATCTAGACAGCA 1960
DB 1801 TCTACAGAGATTCATAGCTAAATAGAAAGTTTGGTGAAGATCTAGACAGCA 1860
QY 1961 GCTGAAATATAAGCACTTTCCTTTGAAACGAACTCCTTGTATTATTCAGCAGTCTGCTC 2020
DB 1861 GCTGAAATATAAGCACTTTCCTTTGAAACGAACTCCTTGTATTATTCAGCAGTCTGCTC 1920
QY 2021 AATATCGGCTCTATATATCTGTTATCTCAATAA 2056
DB 1921 AATATCGGCTCTATATATCTGTTATCTCAATAA 1956

RESULT 6

ADD42825

ID ADD42825 standard; DNA; 1956 BP.

AC ADD42825;

DT 15-JAN-2004 (first entry)

XX Chlamydia pneumoniae antigen polynucleotide SEQ ID NO:63.

XX Chlamydia pneumoniae; Chlamydia; antibiotic; anti-inflammatory;
XX anti-infective; cardiant; anti-atherosclerotic; ophthalmological;
XX vaccine; gene therapy; immune response; pelvic inflammatory disease;
XX tubal obstruction; infertility; male infertility; ocular infection;
XX blindness; acute respiratory tract infection; atherosclerosis;
XX coronary heart disease; gene; ds.

XX Chlamydia pneumoniae.

XX WO2003041560-A2.

XX 22-MAY-2003.

XX 05-NOV-2002; 2002WO-US035624.

XX 06-NOV-2001; 2001US-00012256.

XX 05-DEC-2001; 2001US-00007693.

XX 15-JUL-2002; 2002US-00197220.

XX (CORI-) CORIXA CORP.

XX Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;

PI Probst P;

XX WPI; 2003-441771/41.

XX P-PSDB; ADD42665.

XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful
XX for diagnosing or treating Chlamydial infections, particularly as
XX vaccines for treating or preventing Chlamydial infections, e.g. pelvic
XX inflammatory disease.

XX Disclosure; SEQ ID NO 63; 275pp; English.

XX The present invention describes compounds and methods for diagnosing and
XX treating Chlamydial infection. Chlamydia polynucleotide and protein
XX sequences have antibiotic, anti-inflammatory, anti-infective, cardiant,
XX anti-atherosclerotic and ophthalmological activities, and can be used in
XX vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
XX serodiagnoses or methods from the present invention can be used for the
XX diagnosis or treatment of Chlamydial infections, particularly in
XX humans. The polynucleotides, proteins or compositions are particularly
XX useful for stimulating an immune response in a patient, or for
XX stimulating and/or expanding T cells specific for a Chlamydia protein.
XX Specifically, the polynucleotides, proteins or compositions are useful as
XX vaccines for treating or preventing Chlamydial infections including
XX pelvic inflammatory disease (which results in tubal obstruction and
XX infertility in women), male infertility, ocular infection (which may
XX cause blindness), acute respiratory tract infections, atherosclerosis, or
XX coronary heart disease. The present sequence is used in the
XX exemplification of the present invention.

XX Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

Query Match 90.7%; Score 1956; DB 9; Length 1956;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGTTAATCCTATTGGTCCAGGTCCTATAGACGAAACAGACGACCTCCGCGAGAT 160

DB 1 ATGTTAATCCTATTGGTCCAGGTCCTATAGACGAAACAGACGACCTCCGCGAGAT 60

QY 161 CTTTCTGCTCAAGGATGGAGGCGAGTGCAGAGAAATAGAGTGGGAGCTCAAGAAATA 220

DB 61 CTTTCTGCTCAAGGATGGAGGCGAGTGCAGAGAAATAGAGTGGGAGCTCAAGAAATA 120

QY 221 GCAGGTGCGGAAGCTAAGCCTAAGAAATCTAAGACGATTTCTAGAGCGATCGAGCATC 280

DB 121 GCAGGTGCGGAAGCTAAGCCTAAGAAATCTAAGACGATTTCTAGAGCGATCGAGCATC 180

QY 281 TTGCGTTCTGCAATGATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGTTCTAGT 340

DB 181 TTGCGTTCTGCAATGATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTGTTCTAGT 240

QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGAGCTCAACGACGAGCGGACCT 400

DB 241 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGAGCTCAACGACGAGCGGACCT 300

QY 401 ACAGCTCTCCACCCAGCTTTGATGATTATAAGACTCAAGCGAAACAGCTTACGATCT 460

DB 301 ACAGCTCTCCACCCAGCTTTGATGATTATAAGACTCAAGCGAAACAGCTTACGATCT 360

QY 461 ATCTTACCTCAACATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 520

DB 361 ATCTTACCTCAACATCAGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 521 GTCACTAATATAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 580

DB 421 GTCACTAATATAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

QY 581 GAAACTAAGAAATGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640

DB 481 GAAACTAAGAAATGCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 641 GCTTCGGATAACCAAGCGATTTCTTGAATCTTTAGGTAAATGATGATGATGATGATG 700

```
Db 541 |||||GCTTCGGATAACCAAGCGATCTTTGAGCTCTTTAGGTAAGCTGCTTCCTTCGACCTCTTA 600
QY 701 CAGGCTGCTCTTCCATCTGTGAGCAACAATTAACAAGCAGCTGAGCTTCTTAAAGAG 760
Db 601 |||||CAGGCTGCTCTTCCATCTGTGAGCAACAATTAACAAGCAGCTGAGCTTCTTAAAGAG 660
QY 761 ATGCAAGATAACCCAGTAGTCCCGAGGAAACGGCTGCAATTGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTAGTCCCGAGGAAACGGCTGCAATTGCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATAT 880
Db 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCAATAT 780
QY 881 TTTTCAGACAGAACGCTAGTGGAGCTGTAGAAATGCTAAATCTAATTAACAGATTAAGC 940
Db 781 TTTTCAGACAGAACGCTAGTGGAGCTGTAGAAATGCTAAATCTAATTAACAGATTAAGC 840
QY 941 AACATAGATTGAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 1000
Db 841 AACATAGATTGAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAG 900
QY 1001 AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 1060
Db 901 AAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAG 960
QY 1061 AAAGATCTTAAATAATCAAACTGCGAGATGTTCTGATGTTCCAAATCCAGGAACATACA 1120
Db 961 AAAGATCTTAAATAATCAAACTGCGAGATGTTCTGATGTTCCAAATCCAGGAACATACA 1020
QY 1121 GTTGGAGGCTTCAAGCAACAGGAAGTAGTATTTGGTAGTATTCGTGTTTCCATGCTGTTA 1180
Db 1021 GTTGGAGGCTTCAAGCAACAGGAAGTAGTATTTGGTAGTATTCGTGTTTCCATGCTGTTA 1080
QY 1181 GATGATGCTGAAATGAGACCGCTTCAATTTGATGTTGCTGGTTTGGTCAAGATGATTCAC 1240
Db 1081 GATGATGCTGAAATGAGACCGCTTCAATTTGATGTTGCTGGTTTGGTCAAGATGATTCAC 1140
QY 1241 ATGTTCAATACGGAATCTCTGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCT 1300
Db 1141 ATGTTCAATACGGAATCTCTGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCT 1200
QY 1301 AGACAGCGAAAGCGCTGGAGATGACAGTGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGACAGCGAAAGCGCTGGAGATGACAGTGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1260
QY 1361 GCTTAGAGCGGCTCTAGTTAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGA 1420
Db 1261 GCTTAGAGCGGCTCTAGTTAAAGCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGTGTTGTGAGCGCAGAGTTCCTCCGCTGCGAGCAAGTTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGTGTTGTGAGCGCAGAGTTCCTCCGCTGCGAGCAAGTTCTATA 1380
QY 1481 GGGTCATCTGTAACAAGCTTTTACAAGCTTCAAAATCTACAGGTTCTGATTATAAACA 1540
Db 1381 GGGTCATCTGTAACAAGCTTTTACAAGCTTCAAAATCTACAGGTTCTGATTATAAACA 1440
QY 1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGTTAGGCGACGA 1500
QY 1601 AATGATGCACTCGTGTATGATATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTT 1660
Db 1501 AATGATGCACTCGTGTATGATATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTT 1560
QY 1661 CCTAGAGCAAGACAGAGCTCGAGGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTG 1720
Db 1561 CCTAGAGCAAGACAGAGCTCGAGGACAGAAAAAACAAGATCAAGCCCTCGCTAGGGTG 1620
QY 1721 ATTTCTGCAATAGCAGACTCTTGGAGATGCTTATAGTCAAGTTTGGGCACTACAATCT 1780
```

```
Db 1621 ATTTCTGGCAATAGCAGAACTCTTGGAGATGTTCTATAGTCAAGTTTCGGCACTACAATCT 1680
QY 1781 GTATTCAGAGATCATCCAGTCGAATCTCTCAAGCGAATAATAGGAGATCAGACAAAAGCTT 1840
Db 1681 GTATTCAGAGATCATCCAGTCGAATCTCTCAAGCGAATAATAGGAGATCAGACAAAAGCTT 1740
QY 1841 ACATCGSCAGTGACAAAAGCTTCCACAGTTTGGGTATCTCTTATGTGCAACTTTCTAATGAC 1900
Db 1741 ACATCGSCAGTGACAAAAGCTTCCACAGTTTGGGTATCTCTTATGTGCAACTTTCTAATGAC 1800
QY 1901 TCTACACAGAAAGTTTATAGCTAAATTAAGAAGTTTGTTCCTGCTGAAGATCTAGGACAGCA 1960
Db 1801 TCTACACAGAAAGTTTATAGCTAAATTAAGAAGTTTGTTCCTGCTGAAGATCTAGGACAGCA 1860
QY 1961 GCTGMAATAAAGACACTTTCCTTTGAAACGAACCTCCTTGTGTTTATTCAGCAGGTCGTCGTC 2020
Db 1861 GCTGMAATAAAGACACTTTCCTTTGAAACGAACCTCCTTGTGTTTATTCAGCAGGTCGTCGTC 1920
QY 2021 AATATCGGCTCTCTATATATCTGTTATCTCTCAATAA 2056
Db 1921 AATATCGGCTCTCTATATATCTGTTATCTCTCAATAA 1956
```

RESULT 7

AAD02064

ID AAD02064 standard; DNA; 1852 BP.

XX AC AAD02064;

XX AC AAD02064;

XX XX 26-MAR-2001 (first entry)

XX XX 5'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.

XX XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;

XX XX upper respiratory tract disease; bronchitis; sinusitis;

XX XX acute respiratory disease; cough; sore throat; hoarseness; fever;

XX XX vaccine; immunisation; treatment; truncation mutant; ds.

XX XX Chlamydia pneumoniae.

OS Synthetic.

XX Key Location/Qualifiers

XX CDS 1..1752

XX FT /*tag= a

XX FT /product= "5'-truncated Chlamydia pneumoniae 76kDa

XX FT protein"

XX FT /transl_except= (pos:1489..1491, aa:Ile)

XX XX MO200066739-A2.

XX XX 09-NOV-2000.

XX XX 03-MAY-2000; 2000WO-CA000511.

XX XX 03-MAY-1999; 99US-0132270P.

XX XX 30-JUN-1999; 99US-0141276P.

XX XX (AVET) AVENTIS PASTEUR LTD.

XX XX Murdin AD, Comen RP, Wang J, Dunn P;

XX XX WPI; 2000-687542/67.

XX XX P-PSDB; AAY71955.

XX XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful

XX XX for vaccinating against Chlamydia infections.

XX XX Claim 2b; Page 97-99; 112pp; English.

XX XX The present sequence is a DNA coding for 5'-truncated Chlamydia

XX XX pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the

XX XX diagnosis, prevention and treatment of C. pneumoniae infections (e.g.

XX XX pneumonia, upper respiratory tract disease, bronchitis, sinusitis and

Signature

CC acute respiratory disease such as cough, sore throat, hoarseness, fever;
CC and abnormal chest sounds on auscultation). C. pneumoniae sequence is
CC also used as vaccines for immunising humans against diseases caused by C.
CC pneumoniae
XX
SQ

Sequence 1852 BP; 578 A; 409 C; 407 G; 458 T; 0 U; 0 Other;
Query Match 85.8%; Score 1848.8; DB 3; Length 1852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 305 ATGACTCTGGCAGATAAGCTGGTATTCGCTCTAGTAACACGCTGCTCTTCTACTAGCAGA 364
Db 1 ATGAGTCTGGCAGATAAGCTGGTATTCGCTCTAGTAACACGCTGCTCTTCTACTAGCAGA 60
QY 365 TCTGCAGACGCTGGATCTCAACACACGACGACCGCACCTACGCTCTCCACCCACGCTTTGAT 424
Db 61 TCTGCAGACGCTGGATCTCAACACACGACGACCGCACCTACGCTCTCCACCCACGCTTTGAT 120
QY 425 GATTATAGACTCAAGCGCAACAGCTTACGATACCTATCTTTTACCTCAACATCACTAGCT 484
Db 121 GATTATAGACTCAAGCGCAACAGCTTACGATACCTATCTTTTACCTCAACATCACTAGCT 180
QY 485 GACATACAGGCTGCTTTGGTGCAGCTCCAGGATGCTGTCACTATATATAAGATACACG 544
Db 181 GACATACAGGCTGCTTTGGTGCAGCTCCAGGATGCTGTCACTATATATAAGATACACG 240
QY 545 GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAATGCCGATGCACTT 604
Db 241 GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAATGCCGATGCACTT 300
QY 605 AAAGTTGGCGCGAAATTTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATCTT 664
Db 301 AAAGTTGGCGCGAAATTTACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGATCTT 360
QY 665 GACTCTTTAGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA 724
Db 361 GACTCTTTAGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA 420
QY 725 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 784
Db 421 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
QY 785 GGGAAACCCCTGCAATTTGCTCAATTTAGTGAATGATGATGATGATGATGATGATGATGAT 844
Db 481 GGGAAACCCCTGCAATTTGCTCAATTTAGTGAATGATGATGATGATGATGATGATGATGAT 540
QY 845 ATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTTCAGGACAGACGCTAGTGGGA 904
Db 541 ATAGAGAAAGATGGAATGCGATTAGGATGCAATATTTTTCAGGACAGACGCTAGTGGGA 600
QY 905 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTTAAAGCAGCA 964
Db 601 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTTAAAGCAGCA 660
QY 965 ATCGCTACTGTAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAAATT 1024
Db 661 ATCCCTACTGTAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAAATT 720
QY 1025 CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTCAGAAAAGATCTTAAAAATATCAAACT 1084
Db 721 CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTCAGAAAAGATCTTAAAAATATCAAACT 780
QY 1085 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACAGGA 1144
Db 781 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGAGGCTCCAAAGCAACAGGA 840
QY 1145 AGTAGTATTTGCTAGTATTCGTTTCCATGCTGTAGATGATGATGATGATGATGATGATGAT 1204
Db 841 AGTAGTATTTGCTAGTATTTCCATGCTGTAGATGATGATGATGATGATGATGATGATGAT 900
QY 1205 TCCATTTTGTATGCTGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATTCCTGAT 1264

Db 901 TCCATTTTGTATGCTGGGTTTCGTCAGATGATTCACATGTTCAATACGAAAAATCCTGAT 960
QY 1265 TCTCAAGCTGCCCAACAGGAGCTCGGACGACCAAGCTAGACAGCGAAGCGCTGGAGAT 1324
Db 961 TCTCAAGCTGCCCAACAGGAGCTCGGACGACCAAGCTAGACAGCGAAGCGCTGGAGAT 1020
QY 1325 GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGGTAAA 1384
Db 1021 GACAGTCTGCTGCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGGTAAA 1080
QY 1385 GCTGGGCAACACAGGCGATCTCAATGCTTTTAGGACAGATGCTGCTGCTGCTGCTGCTG 1444
Db 1081 GCTGGGCAACACAGGCGATCTCAATGCTTTTAGGACAGATGCTGCTGCTGCTGCTGCTG 1140
QY 1445 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCATAGSGTCACTGTGTAACACAGCTTTAC 1504
Db 1141 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCATAGSGTCACTGTGTAACACAGCTTTAC 1200
QY 1505 AAGACCTCAAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1564
Db 1201 AAGACCTCAAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
QY 1565 TACAAATCCATCAATGATGCTATGCTAGGACGCAAAATGATGCGACTCGTGTGATGATA 1624
Db 1261 TACAAATCCATCAATGATGCTATGCTAGGACGCAAAATGATGCGACTCGTGTGATGATA 1320
QY 1625 AACAAATGTAAGTACCCCGCTCTCACAGCATCGCTTCTTAGAGCACGAAACAGAAAGCTCGA 1684
Db 1321 AACAAATGTAAGTACCCCGCTCTCACAGCATCGCTTCTTAGAGCACGAAACAGAAAGCTCGA 1380
QY 1685 GGACCAAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGSCAATAGCAGAACTCTT 1744
Db 1381 GGACCAAGAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGSCAATAGCAGAACTCTT 1440
QY 1745 GGAGATCTCTATAGTCAAGTTTCGGGACTACATCTGTANTGAGATCACTCCAGTCCAAAT 1804
Db 1441 GGAGATCTCTATAGTCAAGTTTCGGGACTACATCTGTANTGAGATCACTCCAGTCCAAAT 1500
QY 1805 CCTCAACGCAATATAGAGGATCAGACAAAAGCTTACATCGGCGAGTGACAAAAGCCCTCCA 1864
Db 1501 CCTCAACGCAATATAGAGGATCAGACAAAAGCTTACATCGGCGAGTGACAAAAGCCCTCCA 1560
QY 1865 CAGTTTGGCTATCTTANGTGCAACTTTCTAATGACTCTACACAGAAAGTTTCAATAGCTAAA 1924
Db 1561 CAGTTTGGCTATCTTANGTGCAACTTTCTAATGACTCTACACAGAAAGTTTCAATAGCTAAA 1620
QY 1925 TTAGAAAGTTTGTGCTGAAGGATCTAGCAGCAGCTGAAATATAAGCACTTTCCCTTT 1984
Db 1621 TTAGAAAGTTTGTGCTGAAGGATCTAGCAGCAGCTGAAATATAAGCACTTTCCCTTT 1680
QY 1985 GAAACGAACTCCTTTGTTTATTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCGTT 2044
Db 1681 GAAACGAACTCCTTTGTTTATTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCGTT 1740
QY 2045 TATCTCCAAATAACACACCTTAAGTTTCGTTTGGAGAGATATATGCTGCTTGGTAAAG 2104
Db 1741 TATCTCCAAATAACACACCTTAAGTTTCGTTTGGAGAGATATATGCTGCTTGGTAAAG 1800
QY 2105 CCCTTTGTTGAGGCTTACCAACACACCTAGAACGATCTTCAATAATAAATAAAGA 2156
Db 1801 CCCTTTGTTGAGGCTTACCAACACACCTAGAACGATCTTCAATAATAAATAAAGA 1852

RESULT 8

AAD02066

ID AAD02066 standard; DNA; 2238 BP.

XX AAD02066;

AC AAD02066;

XX 26-MAR-2001 (first entry)

XX

XX C. pneumoniae 76 kDa protein truncation mutant fusion gene.

76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
upper respiratory tract disease; bronchitis; sinusitis;
acute respiratory disease; cough; sore throat; hoarseness; fever;
vaccine; immunisation; treatment; truncation mutant; fusion gene; ds.
Chlamydia pneumoniae.
Synthetic.

Key Location/Qualifiers
misc_feature 1..665
/tag= a
/note= "This part of the sequence is unrelated to C.
pneumoniae 76 kDa gene"
CDS 766..2238
/tag= b
/product= "Truncated Chlamydia pneumoniae 76kDa protein"
misc_feature 2122..2238
/tag= c
/note= "This part of the sequence is unrelated to C.
pneumoniae 76 kDa gene"

WO20066739-A2.
09-NOV-2000.
03-MAY-2000; 2000WO-CA000511.
03-MAY-1999; 99US-0132270P.
30-JUN-1999; 99US-0141276P.
(AVET) AVENTIS PASTEUR LTD.
Murdin AD, Oomen RP, Wang J, Dunn P;
WPI; 2000-687542/67.
P-PSDB; MAY71957.

Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.
Claim 32; Fig 3; 112pp; English.

The present sequence is a DNA coding for a fusion protein comprising a
truncated Chlamydia pneumoniae 76 kDa protein and vector-encoded
residues. C. pneumoniae 76 kDa protein is used in the diagnosis,
prevention and treatment of C. pneumoniae infections (e.g. pneumonia,
upper respiratory tract disease, bronchitis, sinusitis and acute
respiratory disease such as cough, sore throat, hoarseness, fever; and
abnormal chest sounds on auscultation). C. pneumoniae sequence is also
used as vaccines for immunising humans against diseases caused by C.
pneumoniae

Sequence 2238 BP; 689 A; 459 C; 497 G; 593 T; 0 U; 0 Other;
Query Match 68.0%; Score 1465.8; DB 3; Length 2238;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1478; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 60
666 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 725
61 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 120
726 ATAAATCTTAAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 785
121 AGGCTCTATAGCAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 180
786 AGGCTCTATAGCAACACAGGCTCCGATTAATATTAGTGAGAGCTTTTATTTTATTTT 845
181 GCGAGTGCGCAATTAAGAGTGCGGAGCTCAAGAGTACGAGTCCGGAAGCTAAGCC 240
846 GCGAGTGCGCAATTAAGAGTGCGGAGCTCAAGAGTACGAGTCCGGAAGCTAAGCC 905

241 TAAAGATCTAAGACCGATTTCTGTAGAGCGATGAGCATCTTGGTCTTGCAGTGAATGC 300
906 TAAAGATCTAAGACCGATTTCTGTAGAGCGATGAGCATCTTGGTCTTGCAGTGAATGC 965
301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTAG 360
966 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTAG 1025
361 CAGATCTCGAGAGTGGACTCAACGACGAGCGGACCTAGCGCTCTCCACCGAGTT 420
1026 CAGATCTCGAGAGTGGACTCAACGACGAGCGGACCTAGCGCTCTCCACCGAGTT 1085
421 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACT 480
1086 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTACCTCAACATCACT 1145
481 AGCTGACATACAGGCTCTTTGGTGGAGCTCCAGGATGCTGTCACTAATATAAGGATAC 540
1146 AGCTGACATACAGGCTCTTTGGTGGAGCTCCAGGATGCTGTCACTAATATAAGGATAC 1205
541 AGCGCTACTGATGAGCAACCGCAATCGCTGGAGTGGGAACTAAGATGCCGATGC 600
1206 AGCGCTACTGATGAGCAACCGCAATCGCTGGAGTGGGAACTAAGATGCCGATGC 1265
601 AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 660
1266 AGTTAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGAT 1325
661 TCTTGACTCTTTAGGTAAACTGACTCTTCGACCTCTTACAGGCTCTCTTCCCAATC 720
1326 TCTTGACTCTTTAGGTAAACTGACTCTTCGACCTCTTACAGGCTCTCTTCCCAATC 1385
721 TGTAGCAACAATAAACAAGCAGCTGAGCTCTTAAAGAGATCAAGATTAACCAAGTAGT 780
1386 TGTAGCAACAATAAACAAGCAGCTGAGCTCTTAAAGAGATCAAGATTAACCAAGTAGT 1445
781 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
1446 CCCAGGGAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 1505
841 ACAGATAGAGAAAGATGGAAATCGAATTAGGATGCAATATTTTCAGGACAGAAACGCTAG 900
1506 ACAGATAGAGAAAGATGGAAATCGAATTAGGATGCAATATTTTCAGGACAGAAACGCTAG 1565
901 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATGATTTCAGTAAAGC 960
1566 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATGATTTCAGTAAAGC 1625
961 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1020
1626 AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCC 1685
1021 AATTCTTCAAGAGCGGAAACAAATGTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
1686 AATTCTTCAAGAGCGGAAACAAATGTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1745
1081 ACCTGACAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1140
1746 ACCTGACAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACA 1805
1141 AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC 1200
1806 AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGAC 1865
1201 CGCTTCCATTTGATGCTCTGGGTTTCTGTCAGATGATTCATGTTCAATACGGAATAATCC 1260
1866 CGCTTCCATTTGATGCTCTGGGTTTCTGTCAGATGATTCATGTTCAATACGGAATAATCC 1325
1261 TGATTTCTCAAGCTGCGCAACAGGAGCTCGCAGCACAAGCTAGACGACGGAAGCCCTGG 1320
1926 TGATTTCTCAAGCTGCGCAACAGGAGCTCGCAGCACAAGCTAGACGACGGAAGCCCTGG 1985

QY 1321 AGATGACAGTGTCTGCGAGCGTGGCAGATGCTCAGAAGCTTTAGAGCGGCTCTAGG 1380
 Db 1986 AGATGACAGTGTCTGCGAGCGTGGCAGATGCTCAGAAGCTTTAGAGCGGCTCTAGG 2045
 QY 1381 TAAAGCTGGGCAACAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 1440
 Db 2046 TAAAGCTGGGCAACAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 2105
 QY 1441 TGTGAGGCGAGGAGTCTTCCGCGTGGCAGCAAGTCTCTATAG 1481
 Db 2106 TGTGAGGCGAGGAG-TACTCCGCGTGGCAGCAAGTCTCTATAG 2145

RESULT 9
 AAD02065
 ID AAD02065 standard; DNA; 1456 BP.
 XX
 AC AAD02065;
 DT 26-MAR-2001 (first entry)
 XX
 DE 3'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.
 XX
 KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; truncation mutant; ds.
 XX
 OS Chlamydia pneumoniae.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 101..1456
 FT /*tag= a
 FT /product= "3'-truncated Chlamydia pneumoniae 76kDa
 FT /note= "The coding region does not include stop codon"
 FT /partial
 XX
 XX W0200066739-A2.
 XX
 XX 09-NOV-2000.
 XX
 XX 03-MAY-2000; 2000WO-CA000511.
 XX
 XX 03-MAY-1999; 99US-0132270P.
 XX
 XX 30-JUN-1999; 99US-0141276P.
 XX
 XX (AVET) AVENTIS PASTEUR LTD.
 XX
 XX Murdin AD, Oomen RP, Wang J, Dunn P;
 XX
 XX WFI; 2000-687542/67.
 XX
 XX P-PSDB; AAY71956.
 XX
 XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 XX for vaccinating against Chlamydia infections.
 XX
 XX Claim 2c; Page 102-104; 112pp; English.
 XX
 XX The present sequence is a DNA coding for 3'-truncated Chlamydia
 XX pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the
 XX diagnosis, prevention and treatment of C. pneumoniae infections (e.g.
 XX pneumonia, upper respiratory tract disease, bronchitis, sinusitis and
 XX acute respiratory disease such as cough, sore throat, hoarseness, fever;
 XX and abnormal chest sounds on auscultation). C. pneumoniae sequence is
 XX also used as vaccines for immunising humans against diseases caused by C.
 XX pneumoniae
 XX
 XX Sequence 1456 BP; 452 A; 308 C; 331 G; 365 T; 0 U; 0 Other;
 XX
 XX Query Match 67.5%; Score 1455; DB 3; Length 1456;
 XX Best Local Similarity 100.0%; Pred. No. 0;

Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATAAATCTTTTAAACAGGCTCGCATTAATTAATAGTGAAGCTTTTATTTTATTTT 60
 Db 1 AATAAATCTTTTAAACAGGCTCGCATTAATTAATAGTGAAGCTTTTATTTTATTTT 60
 QY 61 AATAAATAAATAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
 Db 61 AATAAATAAATAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
 QY 121 AGGTCTTATAGACGAAACGAAACGACCTCCGAGATCTTTCTGCTCAAGGATTTGA 180
 Db 121 AGGTCTTATAGACGAAACGAAACGACCTCCGAGATCTTTCTGCTCAAGGATTTGA 180
 QY 181 GCGGAGTGCAGCAAAATAAGAGTGGGAGCTCAAGAATAGCAGGTGCGGAAGCTTAAGCC 240
 Db 181 GCGGAGTGCAGCAAAATAAGAGTGGGAGCTCAAGAATAGCAGGTGCGGAAGCTTAAGCC 240
 QY 241 TAAAGAATCTAAGACCGGATTTCTGTAGAGCGATGGAGCATCTTGGGTTCTGCAGTGAATGC 300
 Db 241 TAAAGAATCTAAGACCGGATTTCTGTAGAGCGATGGAGCATCTTGGGTTCTGCAGTGAATGC 300
 QY 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTTCTTCTAGTAAACAGCTCGCTTCTACTAG 360
 Db 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTTCTTCTAGTAAACAGCTCGCTTCTACTAG 360
 QY 361 CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACCCTCGGCTCTCCACCCACGTT 420
 Db 361 CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACCCTCGGCTCTCCACCCACGTT 420
 QY 421 TGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATCTATCTTTTACCTCAACATCACT 480
 Db 421 TGATGATTATAAGACTCAAGCGCAAAACAGCTTACGATCTATCTTTTACCTCAACATCACT 480
 QY 481 AGCTGACATACAGGCTGCTTGGTGGAGCTCCAGAGTCTGCTCACTAATAATAAGATAC 540
 Db 481 AGCTGACATACAGGCTGCTTGGTGGAGCTCCAGAGTCTGCTCACTAATAATAAGATAC 540
 QY 541 AGCGGCTACTGATGAGAAACCGCAATCGCTGGAGTGGGAAACTAAGAATGCCGATGC 600
 Db 541 AGCGGCTACTGATGAGAAACCGCAATCGCTGGAGTGGGAAACTAAGAATGCCGATGC 600
 QY 601 AGTTAAAGTTGGCGCGCAAAATTAAGAAATAGGAAATATGCTTCGGATAACCAAGCGAT 660
 Db 601 AGTTAAAGTTGGCGCGCAAAATTAAGAAATAGGAAATATGCTTCGGATAACCAAGCGAT 660
 QY 661 TCTTGACTCTTTAGTAAACTGACTTCTTCGACCTTACAGGCTGCTCTTCTCCAATC 720
 Db 661 TCTTGACTCTTTAGTAAACTGACTTCTTCGACCTTACAGGCTGCTCTTCTCCAATC 720
 QY 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
 Db 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
 QY 781 CCAGGGAACCCCTGCAATTTGCTCAATTTAGTTGATCAGACAGATGCTACAGGAC 840
 Db 781 CCAGGGAACCCCTGCAATTTGCTCAATTTAGTTGATCAGACAGATGCTACAGGAC 840
 QY 841 ACAGATAGAGAAAGATGGAATCGGATAGGATGCATATTTTTCAGGACAGAAACGCTAG 900
 Db 841 ACAGATAGAGAAAGATGGAATCGGATAGGATGCATATTTTTCAGGACAGAAACGCTAG 900
 QY 901 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATAGATTAGCTAAAGC 960
 Db 901 TGGAGCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATAGATTAGCTAAAGC 960
 QY 961 AGCAATCGCTACTGCTAAGACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
 Db 961 AGCAATCGCTACTGCTAAGACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
 QY 1021 AATTTCTCAAGAAGCGGACAAATGTTATACAGGCTGAGAAAAGATCTTAAAAATATCAA 1080
 Db 1021 AATTTCTCAAGAAGCGGACAAATGTTATACAGGCTGAGAAAAGATCTTAAAAATATCAA 1080

```
QY 1081 ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAATACAGTGTGGAGGCTCCAGCAACA 1140
Db 1081 ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAATACAGTGTGGAGGCTCCAGCAACA 1140
QY 1141 AGGAAGTAGATTGGTAGTATCGTTTCCATGCTGTAGATGATGCTGCAAAATGAGAC 1200
Db 1141 AGGAAGTAGATTGGTAGTATCGTTTCCATGCTGTAGATGATGCTGCAAAATGAGAC 1200
QY 1201 CGCTTCCATTTTGTATGTCGGTTTCTGATGATTCAGATGATTCAGATGTTCAATACGGAATCC 1260
Db 1201 CGCTTCCATTTTGTATGTCGGTTTCTGATGATTCAGATGATTCAGATGTTCAATACGGAATCC 1260
QY 1261 TGAATCTCAAGTCCCAACAGGAGCTCGCAGCAAAAGCTTGAAGCGGCTTGG 1320
Db 1261 TGAATCTCAAGTCCCAACAGGAGCTCGCAGCAAAAGCTTGAAGCGGCTTGG 1320
QY 1321 AGATGACAGTGTCTGCTGAGGCTGGCAGATGCTCAGAAAGCTTTAGAAGCGGCTTAGG 1380
Db 1321 AGATGACAGTGTCTGCTGAGGCTGGCAGATGCTCAGAAAGCTTTAGAAGCGGCTTAGG 1380
QY 1381 TAAAGCTGGGCAACAAACAGGCTACTCAATGCTTTAGGACAGATGCTTCTGCTGCTGT 1440
Db 1381 TAAAGCTGGGCAACAAACAGGCTACTCAATGCTTTAGGACAGATGCTTCTGCTGCTGT 1440
QY 1441 TGTGAGCGCAGGAGT 1455
Db 1441 TGTGAGCGCAGGAGT 1455

RESULT 10
AAZ01425_07
Continuation (8 of 11) of AAZ01425 from base 700001 (Complete genome sequence of Chlamy
WP Sequence split into 11 fragments LOCUS AAZ01425 Accession AAZ01425
WP Fragment Name Begin End
WP AAZ01425_01 1 110000
WP AAZ01425_02 100001 210000
WP AAZ01425_03 200001 310000
WP AAZ01425_04 300001 410000
WP AAZ01425_05 400001 510000
WP AAZ01425_06 500001 610000
WP AAZ01425_07 600001 710000
WP AAZ01425_08 700001 810000
WP AAZ01425_09 800001 910000
WP AAZ01425_10 900001 1010000
WP AAZ01425_11 1000001 1038602

Query Match 8.7%; Score 187.6; DB 2; Length 110000;
Best Local Similarity 53.0%; Pred. No. 9.6e-36;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 1167 TTTCATGCTGTGTAGATGATGCTGAAATGAGACCGCTTCCATTTTGTATGCTGGGTTTC 1226
Db 105378 TTTCCTTTGTTGTTGATGATGTAGCAATGAAATGGCAGGATGCAATGCAAGGTTTTC 105437
QY 1227 GTGATGATTCATGTTCAATACGGAATCTGATTTCTCAAGCTGCCAACAGGAGC 1286
Db 105438 GATCTATGATCGAACATTTAATGTAACCAATCTGCAACAGCTAAAGAGCTACAAGCTA 105497
QY 1287 TCCGACACAGCTAGACGCGGCTTGAAGCTGTGAGATGACGTGCTGCGAGCGCTGG 1346
Db 105498 TGGAGGCTCAGCTGA---CTGCCATGTGATCAATCAATGTTGTCGGATGCGGAGCTCC 105554
QY 1347 CAGATGCTCAGAAAGCTTTAGAAGCGGCTCTAGTAAAGCTGGGCAACAAC----- 1397
Db 105555 CAGCGAAATACAGCAATCAAGATGCTCTTGGCAAGCTTTGAACACACCATCAGCAG 105614
QY 1398 AGGCATCTCAATGCTTTTAGGACAGATCGCTTCTGCTGCTGTGTGTGAGCGCAGGAGTTC 1457
Db 105615 ATGGTTTAGCTACAGCTATGCGGCAAGTGGCTTTTTCAGCTGCGCAAGGTTTGGAGGAGCT 105674
QY 1458 CTCGCTGCGACGAAGTTCTATAGGTCATCTGTAACACAGCTTACAAGACCTCAAAAT 1517
```

```
Db 105675 CCGCAGGAACAGCTGGCACTGTCCAGATGAATGTAACACAGCTTTTACAGACAGCGTTTT 105734
QY 1518 CTACAGGTTCT-----GATTATAAACACAGATATACAGAGGTTATGATGCTTCAAAAT 1571
Db 105735 CTTTCGACTTCTTCAAGCTCTTATGTCAGCAGCACTTTCCGATGATATCTGCTTACAAA 105794
QY 1572 CCATCAATGATGCTATGTTAGGGCACGAATGATGCGACTCTGATGTGTGATAAACAAATG 1631
Db 105795 CACTGAACTCTTTATATTCGGAAGCAGAAAGCGGCTGCGATGATGATTTAGTCAAACTG 105854
QY 1632 TAAGTACCCCGCTCTCAACGATCCGTTCTTAGACACGAAACAGAGCTTCGAGGACCAG 1691
Db 105855 CAAATCCCGCTTTTCCAGAAAGCGTTTCTCGTTCTGGCATAGAAAGCTCAAGGACGCGATG 105914
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGCGAATAGCAGAACTCTTTGGAGATG 1751
Db 105915 CAGATCGTACCCNAGAGCAGCGAAACTATTCTCAGATAGCCAAAGCTTAGGTGATG 105974
QY 1752 TCTATAGTCAAGTTTCGGCACTACAATCTGTATGATGATCATCCAGTGAATCCCTCAAG 1811
Db 105975 TATATAGCGCTTACAGGTTCTGGATTCTTTGATGCTACGATTTGAGCAATCCGCAAG 106034
QY 1812 CGAATAATGAGGATTCAGACAAAAGCTTACATCGGCGAGTGACAAGCCTCCACAGTTTG 1871
Db 106035 CAAATCAAGAAGAGATTATGCGAGAGCTCAGGCACTCTATTAGCAAAAGCTCCCACAATTG 106094
QY 1872 GCTATCCTTATGTCACACTTTCTTAATGACTCTACACAGAGTTCAATGATTAATAGAA 1931
Db 106095 GGTATCCTGCTGTTCAAAATCTGTGATGATGTTGCGAAGTTTGTGCAAAATGGAA 106154
QY 1932 GTTTGTTGCTGAAGATCTAGACAGAGCTGAATAAAGCACTTCTCTTTGAAACGA 1991
Db 106155 GAGAGTTTGTGATGGGAACGTTAGTCTCGCAATCTCAAGAGATGCGTTTAGAAAC 106214
QY 1992 ACTCCTTTGTTTATTCAGAGGTTGCTGCTCAATACGCTCTCTATATTCGTGTTATCTCC 2051
Db 106215 AGCCCGCTTTCATTCACAGGTTGTTGTTAAACATCTCTCTATTTCTCTGTTATCTTT 106274
QY 2052 AATAAC 2057
Db 106275 CTTAAC 106280

RESULT 11
AAZ01425_08
Continuation (9 of 11) of AAZ01425 from base 800001 (Complete genome sequence of Chlamy
WP Sequence split into 11 fragments LOCUS AAZ01425 Accession AAZ01425
WP Fragment Name Begin End
WP AAZ01425_00 1 110000
WP AAZ01425_01 100001 210000
WP AAZ01425_02 200001 310000
WP AAZ01425_03 300001 410000
WP AAZ01425_04 400001 510000
WP AAZ01425_05 500001 610000
WP AAZ01425_06 600001 710000
WP AAZ01425_07 700001 810000
WP AAZ01425_08 800001 910000
WP AAZ01425_09 900001 1010000
WP AAZ01425_10 1000001 1038602

Query Match 8.7%; Score 187.6; DB 2; Length 110000;
Best Local Similarity 53.0%; Pred. No. 9.6e-36;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 1167 TTTCATGCTGTGTAGATGATGCTGAAATGAGACCGCTTCCATTTTGTATGCTGGGTTTC 1226
Db 5378 TTTCCTTTGTTGTTGATGATGTAGCAATGAAATGGCAGGATGCAATGCAAGGTTTTC 5437
QY 1227 GTGATGATTCATGTTCAATACGGAATCTGATTTCTCAAGCTGCCAACAGGAGC 1286
Db 5438 GATCTATGATCGAACATTTAATGTAACCAATCTTCTGCAACAGCTAAAGAGCTACAAGCTA 5497
QY 1287 TCCGACACAGCTAGACGCGGCTTGAAGCTGTGAGATGACAGTGTGCTGCGAGCGCTGG 1346
```

Db 5498 TGGAGCTCAGCTGA---CTGCGATGTCAGATCAACTGGTGGTGGGAGCTCC 5554
QY 1347 CAGATGCTGAGAAGCTTTAGACCGGCTCTAGTTAAAGCTGGGCAAC----- 1397
Db 5555 CAGCCGAATACAAAGCAATCAAGATGCTCTTGGCAAGCTTTGAAACCAACCATCAGCAG 5614
QY 1398 AGGGCATCTCAATGCTTTAGGACAGATCGTCTCTGCTGCTGTGTGAGCGCAGGAGTTC 1457
Db 5615 ATGGTTAGCTACAGCTATGGGCAAGTGGCTTTTGGAGCTGCCAAGTTGGAGGCT 5674
QY 1458 CTCGCGTGCAGCAAGTTCTATAGGTCATCTGTAAACAGCTTTCAAGACCTCAAAAT 1517
Db 5675 CCGCAGCAAGCTGGCACTGTCCAGATGAATGTAAACAGCTTTCAACACAGCGTTT 5734
QY 1518 CTACAGTTCT-----GATTATAAACACAGATATCAGCAGTTATGATGCTTACAAAT 1571
Db 5735 CTTGCACTCTTCAAGCTCTTATGACGACGACTTTCCGATGATATCTGCTTACAAA 5794
QY 1572 CCATCAATGATCGCTATGTTAGGCGCAAAATGATGGACTCGTGTGATGATATAAACAATG 1631
Db 5795 CACTGAATCTTATATTCGAAAGCAGAGCGGTCAGTCACTATTAGTCAACTG 5854
QY 1632 TAAGTACCCCGCTCTCACAGTCCGTTCTAGACGACGAAACAGAGCTCGAGGACGAG 1691
Db 5855 CAATCCCGGCTTTCCAGAAGCGTTTCTCGTCTGGCATAGAAAGTCAAGGACGCAAGT 5914
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTTGGAGATG 1751
Db 5915 CAGATGCTAGCCAAAGACGACGAAACTATGTACAGATAGCCAAACGTTAGTGATG 5974
QY 1752 TCTATAGTCAAGTTTCGSCACTACAATCTGTAAATSCAGATCATCCAGTCAATCTCAAG 1811
Db 5975 TATATAGCGGTTACAGGTTCTGATTTCTTGTCTAGATGTGAGCAATCCGCAAG 6034
QY 1812 CGAATATAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAACCTCCACAGTTTG 1871
Db 6035 CAATCAAGAAGAGATATGACAGAGCTCAGCGATCTATTAGCAAGCTCCACAAATTG 6094
QY 1872 GCTATCTTATGTCAACTTTCTAATGACTCTACACAGAACTTCATAGTAAATTAGAAA 1931
Db 6095 GGTATCTGCTGTTGAGAACTTCTGTGATAGCTGTCGAGAGTTTGTGTCACAAATGAAA 6154
QY 1932 GTTTGTTGCTGAAGATCTAGGACAGCAGCTGAAATATAAAGCACTTTCCTTTGAAACGA 1991
Db 6155 GAGAGTTTGTGATGGGAACTAGTCTCGCAGAACTCTCAAGAGAACTGCGTTTGAAGAAC 6214
QY 1992 ACTCTGTTTATTCAGAGTGTGTCATATCGGCTCTCTATATCTTATGTTATCTCC 2051
Db 6215 AGCCGCTTTCAATTCACAGGTGTGGTAAACATTTGTTCTTCTATCTCTGTTATCTTT 6274
QY 2052 AATAAC 2057
Db 6275 CTTAAC 6280

RESULT 12

AAA63292

ID AAA63292 standard; DNA; 1537 BP.

XX

AC AAA63292;

XX

DT 02-FEB-2001 (first entry)

XX

DE C. trachomatis LGV II clone 12-G3-83 coding sequence.

XX

KW Chlamydial infection; sexually transmitted disease;

KW pelvic inflammatory disease; PID; tubal obstruction; infertility;

KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;

KW coronary heart disease; antibacterial; ss.

XX

OS Chlamydia trachomatis.

XX

PN WO200034483-A2.
XX 15-JUN-2000.
XX 08-DEC-1999; 99WO-US029012.
XX 08-DEC-1998; 98US-00208277.
XX 08-APR-1999; 99US-00288594.
PR 01-OCT-1999; 99US-00410568.
XX 22-OCT-1999; 99US-00426571.
XX (CORI-) CORIXA CORP.
PA
XX
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises amino
PT acid sequence encoded by polynucleotide sequence.
XX
XX Claim 1; Page 128-129; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamydiae are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC infection is one of the most common sexually transmitted diseases and can
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is a
CC the leading cause of preventable blindness worldwide. C. pneumonia is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a nucleic acid sequence
CC isolated in the present invention
XX
SQ Sequence 1537 BP; 459 A; 307 C; 357 G; 414 T; 0 U; 0 Other;

Query Match 8.7%; Score 187.2; DB 3; Length 1537;

Best Local Similarity 53.0%; Pred. No. 1.9e-36;

Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 1167 TTTCCATGCTTTAGATGATGCTGAAAATAGACCGCTCCATTGATGCTCGGTTTC 1226

Db 186 TTTCTTGTGCTTGTATGATGATGACAAATGAAATGGCAGGATTGCAATCAAGTTTC 245

QY 1227 GTCAATGATTCACATGTTCAATACGGAATCTGATTCCTCAAGCTGCCAACAGGAGC 1286

Db 246 GATCTATGATCGAACAAATTTAATGTAACAAATCTTGAACAGCTTAAGAGCTAAGCTA 305

QY 1287 TCGCAGCACAAAGCTAGACGACGAAAGCCCTCGAGATGACAGTGTCTGACGCGTGG 1346

Db 306 TGGAGGCTCAGTGA---CTGCCATGTCAGATCAACTGGTTGTGCGGATGGCGAGTCC 362

QY 1347 CAGATGCTCAGAAAGCTTTAGAACCGCTCTAGTAAAGCTGGGCAACAC----- 1397

Db 363 CAGCCGAAATACAAAGCAATCAAAGATGCTCTTCGCAAGCTTTGAAACACCAATCAGCAG 422

QY 1398 AGGGCATCTCAATGCTTTAGGACAGATCGTCTTCTGCTGTTGTGAGCGCAGGAGTTC 1457

Db 423 ATGGTTTAGCTACAGCTATGGGCAAGTGGCTTTTGCAGCTGCCAAGTTGGAGGAGCT 482

QY 1458 CTCGCGCTGCAGCAAGTTCTATAGGCTCATCTGTAAAAAGCTTTTACAGACCTCAAAAT 1517

Db 483 CCGCAGCAACAGCTGGCACTGTCCAGATGAATGTAACAGCTTTTACAGACAGCGTTT 542

QY 1518 CTACAGTTCT-----GATTATAAACACAGATATCAGCAGGTTATGCTTACAAAT 1571

Db 543 CTTGCACTTCTCCAGCTCTTATGACGACGACACTTTCCGATGGATATCTCTGTTTCAAAA 602

QY 1572 CCATCAATGATGCTTATGGTAGGCGCAAGATGATGCGACTCGTGATGTATAACAATG 1631

```

603 CACTGAACCTCTTTATATCCGAAAGCAGAGCGGCTGCACTAGCTATTAGTCAAACAGT 662
1632 TAAGTACCCCGCTCTACAGATCCGTTCTTAGAGCAGACAGAGCTGAGACAG 1691
663 CAAATCCCGCGCTTCCAGAGCGTTCTCGTTCTGGCATAGAAAGTCAAGGACGAGT 722
1692 AAAAAACAGATCAAGCCCTCTAGGCTGATTCTGGCAATAGCAGAACTCTTGGAGATG 1751
723 CAGATGCTAGCCAAAGACAGCAGAACTATTGTTCAGAGATGCAAACTGTAGTGATG 782
1752 TCTATAGTCAAGTTGGGCTACATACTGTGAATGCAAGATCATCCAGTCGAATCCTCAAG 1811
783 TATATAGCGGCTTACAGGTTCTGGATTCTTTGATGTCTACGATTGTGAGCAATCCGCAAG 842
1812 CGAATATAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCCCTCCAGTTTG 1871
843 CAAATCAAGAGAGATATGAGAGCTTCAGGATCTATTAGCAAGCTCCCAATTTG 902
1872 GCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAACTTCATAGCTAAATAGAAA 1931
903 GGTATCCTGTGTTTCAGAACTCTGTGGATAGCTTTCGAGAACTTTGCTGCACAAATGGAA 962
1932 GTTTGTTGCTGAAGATCTAGGACAGCAGCTGAAATATAAGAGCTTTCTTTGAAACGA 1991
963 GAGAGTTTGTGTGGGAACTAGTCTCGCAGAACTCTCAAGAAATGCGTTTAGAAAAC 1022
1992 ACTCCTTGTATTACAGAGCTGCTGCTCATATCGGCTCTCTATATCTGTTATCTCC 2051
1023 AGCCCGCTTTCANTCAACAGGTGTGGTAAACATTTGCTTCTATTCTCTGTTATCTTT 1082
2052 AATAAC 2057
1083 CTTAAC 1088

RESULT 13
AAH56196
ID AAH56196 standard; DNA; 1537 BP.
XX
AC AAH56196;
XX
DT 05-SEP-2001 (first entry)
XX
DE Chlamydia trachomatis LGV II DNA clone 12-G3-83.
XX
KW Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; Cap1; CT529; OMCB;
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA; ds.
XX
OS Chlamydia trachomatis.
XX
PN W0200140474-A2.
XX
PD 07-JUN-2001.
XX
PF 04-DEC-2000; 2000WO-US032919.
XX
PR 03-DEC-1999; 99US-00454584.
XX
PR 19-APR-2000; 2000US-00556877.
XX
PR 20-JUN-2000; 2000US-00598419.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX
XX WPI; 2001-374831/39.
XX
PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT inflammatory disease, trachoma, acute respiratory tract infections,
PT atherosclerosis and heart disease.
XX

```

```

PS XX
CC The present nucleotide sequence is provided in a specification relating
CC to compounds and methods for the treatment and diagnosis of chlamydial
CC infection. The compounds provided include polypeptides and fusion
CC proteins comprising immunogenic portions of Chlamydia antigens and DNA
CC sequences encoding such polypeptides. They are useful for vaccinating
CC against chlamydial infection, which causes pelvic inflammatory disease,
CC trachoma, acute respiratory tract infections, atherosclerosis and heart
CC disease
XX
SQ Sequence 1537 BP; 459 A; 307 C; 357 G; 414 T; 0 U; 0 Other;

Query Match      8.7%; Score 187.2; DB 4; Length 1537;
Best Local Similarity 53.0%; Pred. No. 1.9e-36;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY 1167 TTTCCATGCTGTAGATGATGCTGAAATAGACCGCTTCCATTTTGTATCTGTGGTTTC 1226
Db 186 TTTCCCTTGTGCTTGTATGATGTAGACAATGAATGGCAGCGATTGCAATCAAGGTTTTC 245
QY 1227 GTCAGATGATTCACATGTTCAATACGGAAATCCTGATTTCAAGCTGCCCAACAGAGC 1286
Db 246 GATCTATGATCGAACAAATTTAATGTAAACAATCCTGCAACAGCTAAAGAGCTACAAGCTA 305
QY 1287 TCGCAGCACAAAGCTAGAGCAGCGAAAGCGCTGGAGATGACAGTGTCTGCGAGCGCTGG 1346
Db 306 TGGAGGCTCAGTGA---CTGCCATGTCAGATCAACTGTTGTTGGATGGCGAGCTCC 362
QY 1347 CAGATGCTCAGAAAGCTTTAGAAGCGCTCTAGTAAAGCTGGGCAACAAC----- 1397
Db 363 CAGCGGAAATACAAAGCAATCAAAGATGCTCTTCCGCAAGCTTTGMAACAACCATCAGCAG 422
QY 1398 AGGCACTACTCAATGCTTTAGGACAGATCGTTCTCTGCTGCTGTGTGTCAGCGAGGCTTC 1457
Db 423 ATGTTTAGCTACAGCTATGGGCAAGTGCTTTTGCACTGCCAAGCTTTGGAGGAGGCT 482
QY 1458 CTCGCGCTGCAGCAAGTTCTATAGGGTCACTCTCTAAACAAGCTTTTCAAGACCTCAAAAT 1517
Db 483 CCGCAGAACAGCTGGCAGCTGTCAGATGAATGTAAACAAGCTTTTCAAGACAGCGTTT 542
QY 1518 CTACAGGTTCT-----GATTATAAAACACAGATATCAGCAGGTTATGATGCTTACAAT 1571
Db 543 CTTCCGACTTCTTCCAGCTCTTATGACAGCACTTTCCGATGGATATTCTGCTTACAATA 602
QY 1572 CCATCAATGATGCTATGTTAGGCGCAGAAATGATGCGACTCGTGTATGATTAACAATG 1631
Db 603 CACTGAACCTCTTTATATTCGAAAGCAGAAAGCGGCTGCAAGTCACTATTAGTCAAACTG 662
QY 1632 TAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAAGAACAGAGCTTCGAGGAC 1691
Db 663 CAAATCCGCGCTTTCAGAGCGTTTCTGTTCTGCGATAGAAAGTCAAGGACGAGT 722
QY 1692 AAAAAACAGATCAAGCCCTCTGTTAGGCTGATTTCTGGCAATAGCAGAACTCTTTGGAGATG 1751
Db 723 CAGATGCTAGCCAAAGACAGCAGAACTATTGTTCAGAGATAGCCAAACGTTAGTGATG 782
QY 1752 TCTATAGTCAAGTTTTCGGCACTACAACTCTGTAATGAGATCATCCAGTCGAATCCTCAAG 1811
Db 783 TATATAGCCGCTTACAGGTTCTGGAATCTTTGATGTCTAGGATGTGAGCAATCCGCAAG 842
QY 1812 CGAATATAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCCCTCCAGCTTTG 1871
Db 843 CAAATCAAGAGAGATTTATGCAAGCTTCAGCACTCTATTAGCAAGCTCCCAATTTG 902
QY 1872 GCTATCCTTATGTGAACTTTCTAATGACTCTTACACAGAGTTTCACTAGTAAATAGAAA 1931
Db 903 GGTATCCTGTGTTTCAGAAATTTCTGTGATAGCTTTGCAAGGTTTGTGCACAATTTGAAA 962
QY 1932 GTTTGTTGCTGAAGATCTAGGACAGCAGCTGAAATATAAGAGCTTTCTTTGAAACGA 1991
Db 963 GAGAGTTTGTGTGGGAACTAGTCTCGCAGAACTCTCAAGAAATGCGTTTAGAAAAC 1022

```

Example 1; Page 138-139; 295pp; English.

[illegible]

PD 19-JUN-2003.
XX 12-DEC-2002; 2002WO-IB005761.
PF 12-DEC-2001; 2001GB-00029732.
PR 06-AUG-2002; 2002GB-00018233.
PR 14-AUG-2002; 2002GB-00018924.
XX (CHIR-) CHIRON SPA.
PA Grandi G, Ratti G;
XX WPI; 2003-532882/50.
XX P-PSDB; ADD43866.
PT New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.
XX Claim 7; SEQ ID NO 162; 164pp; English.
XX The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This polynucleotide sequence
CC represents one of the 131 DNA sequences encoding the C. trachomatis
CC proteins with immunogenic properties of the invention.
XX
XX Sequence 1944 BP; 612 A; 391 C; 445 G; 496 T; 0 U; 0 Other;
Query March 8.68; Score 186; DB 9; Length 1944;
Best Local Similarity 48.58; Pred. No. 4.2e-36;
Matches 720; Conservative 0; Mismatches 725; Indels 39; Gaps 6;
QY 606 AGGTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGATTCTTG 665
DB 467 AGGTATTGATAAGCTCAATCAGTTAGTTAACTTGAAACACAGATCAGACTTTAAAGG 526
QY 666 ACTCTTTAGTAACTGACTTCCTTCGACCTCTTTACAGGCTGCTCTCTCCAACTCTGTAG 725
DB 527 AAACCTTTAAACCAACACAGACTCTGCAGATCAGATTCAGCGGATTAAATAGTCAGTAGAGA 586
QY 726 CAACAAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCAG 785
DB 587 TCACAAAANTCTGCAGATCAAAATTAACAGATCTGGAAGGACAAAACATAAGTTATG 646
QY 786 GGAATAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCCACACAGA 845
DB 647 AAGCTGTTCTCACTAAGCGAGGAGAGTTATCAAGCTTCTTGAAGCGGAAATTAAGT 706
QY 846 TAGAGAAAGTGAATGCAATGAGGATGATATTTTGCAGACAGAACGCTAGTGGAG 905
DB 707 TAGACAAGCTTTTCAGTCTTAITGTGGATCTGGGATCAAGCCAGGCTGCAGTTCTTC 766
QY 906 CTGTAGAAATGCTAAATCTAATAAGATATAAGCAACATAGATTAGCTTAAGCGACGA 965
DB 767 AAGCAGCAAAATATAGCCCAAGTAATATCGAGCCAGCAAGAAATTAATTTGATGCTG 826
QY 966 TCCTACTGTGTAAGACACAAATAGCTGAAGCTTCAGAAAAAGTTCCCGACTCTCCCAATTC 1025
DB 827 CTGAACGAAGGTAACGAGTTAAACACAGAGATACAGGGCTAACGGACTCGCCTTTAG 886

QY 1026 TTCAGAAGCGGAACAAAATGCTTAATACAGCTCAGAAAGATCTTAAATAATCAAACTG 1085
DB 887 TGAATAAAGCTGAGGAGCAGATTAGTCAAGCAAAAAGATATTCAGAGATCAAACT- 945
QY 1086 CAGATGGTTCTGATGTTTCCAAATCCAGGAATACAGTTGGAG-----GTCCTCAAGCAACA 1140
DB 946 --AGTGGTTCGGATAATTCCTATCGTTGGTCCGAGTGGTTCAGCTGCTTCCGAGGAAGTG 1003
QY 1141 AGGAAGTAGTATTGGTAGTAT-----TCGTGTTTCAATGCTCTTTAGATGATG 1187
DB 1004 CGGTAGGAGCGTTGAAATCCTTAACAATTCAGGAAGAATTTCCCTTGTTCCTTTGATGATG 1063
QY 1188 CTGAAAATGAGACCGCTTCCATTTTCATGCTCTGGGTTTCGTGATGATGATTCACATGTTCA 1247
DB 1064 TAGACATGAATGGCAGCGATTGCAATGCAAGTTTTGATCTATGATCGAACAATTTA 1123
QY 1248 ATACGGAATAATCTGATTTCTCAAGCTGCCCAACAGAGGCTCGCAGCAACAAGCTAGAGCAG 1307
DB 1124 ATGTAAACAATCTCTGCAACAGCTAAAGAGCTAACAGCTATGGAGGCTCAGCTGA---CTG 1180
QY 1308 CGAAAGCCGCTGGAGATGACATGCTGCTGCAGCGCTGGCAGATGCTCAGAAAGCTTTAG 1367
DB 1181 CGATGTCAGATCAACTGGTGGTGGGATGGCGAGCTCCAGCGGAAATACAAAGCAATCA 1240
QY 1368 AAGCGCTCTAGGTAAGCTGGCAACAAC-----AGGGCATACTCAATGCTTTAG 1418
DB 1241 AAGATGCTCTTGGCAGCTTTGAAACACCACTCAACAGATGTTTGTAGCTACAGCTATGG 1300
QY 1419 GACAGATCGCTTCTGCTGCTGTGTGTGAGCGCAGGAGTTCTCCCGCTGCAGCAAGTTCTA 1478
DB 1301 GACAGTGGCTTTTTCAGCTGCCAAGCTTGGAGAGGCTCCGAGGAACAGCTGGCAGCTG 1360
QY 1479 TAGGGTCACTGTAAACAGCTTTTACAGACCTCAAAATCTACAGTTCTGA-----TT 1532
DB 1361 TCAGATGAATGTAACACAGCTTTTACAGACAGCGTTTCTTCGACTTCTTCAGGCTCTT 1420
QY 1533 ATAAAAACAGATATCAGCAGGTTTATGATGCTTTACAAATCCATCAATGATGCCATGGTA 1592
DB 1421 ATGCAGCAGCACTTTCGGATGGATATTTCTGCTTACAAAACACTGAACTCTTATATTCG 1480
QY 1593 GGGCAGCAATGATGCGATCTCGTATGTGATATAAACAATGTAAAGTACCCCGCTCTCAAC 1652
DB 1481 AAAGCAAGAACGCGCTGTCAGTCACTGATTTAGTCAAACTGCAAACTCCGCGCTTCCAGAA 1540
QY 1653 GATCCGCTTCCTAGACACGACAGAAAGCTCSAGGACCAAGAAAACAGATCAAGCCCTCG 1712
DB 1541 GCGTTTCTCGTTCGSCATAGAAAGTCAAGGACGCACTGCAATGCTAGCCAAAGAGCAG 1600
QY 1713 CTAGGGTGAATTTCTGCAATAGCAGAACTCTTTGGAGATGCTTATAGTCAAGTTTTCGGCAC 1772
DB 1601 CAGAAACTATTGTCAGAGATAGCCAAACGTTAGGTGATGATATATAGCCGCTTACAGGTTTC 1660
QY 1773 TACAATCTCTAATGCAATCTCCAGTCAATCTTCAGCGAATATAGAGGATACAGAC 1832
DB 1661 TGATTTCTTTGATGCTACGATTTGAGCAATCCGCAAGTAAATCAAGAAGATTTATGC 1720
QY 1833 AAAAGCTTACATCCGCGAGTGACAAAGCTCCACAGTTTGGCTATCCTTATGTGCACTTT 1892
DB 1721 AGNAGCTCAGGCACTTATTAGCAAGCTCCCAATTTGGGTATCTCTGCTGTTCAAGTT 1780
QY 1893 CTATAGACTCTTACACAGAAGTTCTATAGCTAAATTAAGAAAGTTTGTGTGTAAGGATCTTA 1952
DB 1781 CTGCGGATAGCTTTGAGAAAGTTTGTGCGCAATTTGMAAGAGAGTTTGTGTGAGGGAAC 1840
QY 1953 GGCAGCAGCTGAATAAAGACACTTTCCTTTGAACGAACTCTCTGTTTATTTATTCAGCAGG 2012
DB 1841 TAGTCTCCCAAGTCTCCAGAGAATGGTTTTAGAAAACAGCCCGCTTTTCATTCACAGG 1900
QY 2013 TCGTGTCAATATCGCTCTCTATATTTCTGGTTATCTCCATAA 2056
DB 1901 TGTGTGTAACAATGCTTCTCTATTTCTCTGGTTATCTTTCTTTAA 1944

Wed Mar 24 14:13:14 2004

us-10-608-559-1.rng

Page 18

Search completed: March 23, 2004, 20:02:57
Job time : 840.202 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:15 ; Search time 8378.96 Seconds
(without alignments)
11152.641 Million cell updates/sec

Title: US-10-608-559-1

Perfect score: 2156

Sequence: 1 ataaatctttaaacaagg.....gatcttcaataataaaga 2156

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mt.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2156	100.0	2156	6	AX045131	Sequence
2	2156	100.0	10421	1	AE001654	Chlamydia
3	2156	100.0	12548	1	AE002165	Chlamydia
4	2156	100.0	300066	1	AE017159	Chlamydia
5	2156	100.0	300550	1	AP002547	Chlamydia
6	2143.4	99.4	110000	6	AR310754.08	Continuation (9 of
7	1956	90.7	1956	6	AX338293	Sequence
8	1956	90.7	1956	6	AX349491	Sequence
9	1956	90.7	1956	6	AX361992	Sequence
10	1848.8	85.8	1852	6	AX045133	Sequence
11	1465.8	68.0	2238	6	AX045137	Sequence
12	1458	67.6	2148	1	CHT76XDA	L23921 Chlamydia p
13	1455	67.5	1456	6	AX045135	Sequence
14	395	18.3	304769	1	AE016994	Chlamydia
15	187.2	8.7	1537	6	BD264444	Compounds
16	187.2	8.7	1537	6	BD264444	Compounds
17	187.2	8.7	1537	6	AR308168	Sequence
18	187.2	8.7	1537	6	AR321875	Sequence
19	187.2	8.7	1537	6	AX155814	Sequence
20	187.2	8.7	1537	6	AX361664	Sequence
21	187	8.7	10954	1	AE001333	Chlamydia
22	186	8.6	1944	6	AX786284	Sequence
23	185.4	8.6	1941	6	AX362019	Sequence
24	184	8.5	1171	6	AX338243	Sequence
25	184	8.5	1834	6	AX338268	Sequence
26	183	8.5	1983	6	AX338258	Sequence
27	170	7.9	1550	6	BD43971	Chlamydia
28	170	7.9	1550	6	AR212731	Sequence
29	170	7.9	1550	6	AR428600	Sequence
30	148.4	6.9	150	6	AR144067	Sequence
31	118	5.5	14698	1	AE002357	Chlamydia
32	55.4	2.6	2000	6	AX655393	Sequence
33	54.4	2.5	7218	6	I66494	Sequence
34	52.4	2.4	77835	2	PFMAL13P2.3	Continuation (4 of
35	52.4	2.4	110000	2	PFMAL13.23	Continuation (24 of
36	51.6	2.4	2205	6	AR317736	Sequence
37	51.4	2.4	154160	2	BX571709	Sequence
38	50.6	2.3	423	1	MC049	Sequence
39	49.4	2.3	5098	1	NCU51235	Sequence
40	49	2.3	1990	6	AR268644	Sequence
41	49	2.3	2085	6	AR008322	Sequence
42	49	2.3	2085	6	AR027306	Sequence
43	49	2.3	2085	6	AR035771	Sequence
44	49	2.3	2085	6	AR037631	Sequence
45	49	2.3	2085	6	AR078816	Sequence

ALIGNMENTS

RESULT 1
AX045131
LOCUS AX045131
DEFINITION Sequence 1 from Patent WO0066739.
ACCESSION AX045131
VERSION AX045131.1 GI:11343730
KEYWORDS
SOURCE Chlamydia pneumoniae
ORGANISM Chlamydia pneumoniae
REFERENCE
1 Mardin, A.D., Omen, R.P., Wang, J. and Dunn, P.
i (Chlamydia) antigens and corresponding dna fragments and uses
Thereof
Patent: WO 0066739-A 1 (09-NOV-2000;
JOURNAL

2156 bp DNA linear PAT 24-NOV-2000

Aventis Pasteur Limited (CA)	
FEATURES	Location/Qualifiers
source	1..2156
/organism="Chlamydoiphila pneumoniae"	
/mol_type="unassigned DNA"	
/db_xref="taxon:83558"	
101..2056	
CDS	/note="unnamed protein product"
/codon_start=1	
/transl_table=11	
/protein_id="CAC17329.1"	
/db_xref="GI:11343731"	
/translation="MVNPIGPGIDEITPTPADLSAOGLEASANKSAEORLAGAE AKPKESKTVDSVRSWILRSVAVNLADKLGIASSSSSSSTSRSDVDSITATAPT PPPFDDYKTOAQAYDPIFTSTSLADIQALVSLQDAVTNKKDAATDEETAAEW ETKNADAVKVAQITELAKIASDQALIDSLGKLTFFLQALLOSANNKAEALL KEMQNPVVPKRTPAIAQSLVDQDATTATQDEIKQGNRAIRDAYFAGQNASGAVENAKSN NSINISDAKAAITAKTOIAEAOQKFPDSFILQEAQMV:QAEKDLKNIKPADGSV PNPGTTGGSKQSSSISIRVSMILLDAENETASILMSGFQMHMENTPDSQAA QOELAAQARAAGDDSAALADAOALEALGKAGQOQILNALGOIASAAVVA GVPPAAASSIGSSVKOLYKTSKSDYKTOISAGYDAYKSIDNAYGPRANDRDI NNVSTPALTRSPRARTPRGPKTDQALRVISNSRLIDGVYSQVSAQSVNQIITQ SNPQANNEIRKQLTSATVTPQGPYIVQLRVNSDSTQKFLAKLSLFAEGSRTAABIK ALSFETNSLFIQQVLVNIQSLSYGLQ"	
ORIGIN	
Query Match 100.0%; Score 2156; DB 6; Length 2156;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATAAATCTTTAAACACGGCTCGCATTAATTATTAGTGAAGCTTTTATTTTTT 60
Db	1 ATAAATCTTTAAACACGGCTCGCATTAATTATTAGTGAAGCTTTTATTTTTT 60
Qy	61 ATATAAACTAAAGATTTTTTATTTTTTTTGGTTTTTATGTTATCTATTGTTCC 120
Db	61 ATATAAACTAAAGATTTTTTATTTTTTTTGGTTTTTATGTTATCTATTGTTCC 120
Qy	121 AGGTCCTATAGCAAGAAACAGAACGACACCTCCCGCAGATCTTCTGCTCAAGATTGGA 180
Db	121 AGGTCCTATAGCAAGAAACAGAACGACACCTCCCGCAGATCTTCTGCTCAAGATTGGA 180
Qy	181 GCGGAGTCGAGCAATAAGAGTGGGAGCTCAAGAAATAGCAGTGGGAGCTTAAGCC 240
Db	181 GCGGAGTCGAGCAATAAGAGTGGGAGCTCAAGAAATAGCAGTGGGAGCTTAAGCC 240
Qy	241 TAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGGTTCTGCAAGTGAATGC 300
Db	241 TAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGGTTCTGCAAGTGAATGC 300
Qy	301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTTCTAGTAACAGCTCGTCTTCTACTAG 360
Db	301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTTCTAGTAACAGCTCGTCTTCTACTAG 360
Qy	361 CAGATCTGCACAGTGGACTCAACAGCAGCAGCAGCTACAGCTCGCTCCCAACCCAGTT 420
Db	361 CAGATCTGCACAGTGGACTCAACAGCAGCAGCAGCTACAGCTCGCTCCCAACCCAGTT 420
Qy	421 TGATGATTATTAAGACTCAAGCGCAACAGCTTACGATACACTCTTACCTCAACATCACT 480
Db	421 TGATGATTATTAAGACTCAAGCGCAACAGCTTACGATACACTCTTACCTCAACATCACT 480
Qy	481 AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGCTCACTAATAAAGGATAC 540
Db	481 AGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGCTCACTAATAAAGGATAC 540
Qy	541 AGCGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCGGATGC 600
Db	541 AGCGCTACTGATGAGGAAACCGCAATCGCTGGGAGTGGGAAACTAAGAAATGCGGATGC 600
Qy	601 AGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGAT 660
Db	601 AGTTAAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGAT 660
Qy	661 TCCTGACTCTTTAGTAAACAGTACTCTCTCGAGCTCTTACAGGCTGCTCTTCTCCAAATC 720
Db	661 TCCTGACTCTTTAGTAAACAGTACTCTCTCGAGCTCTTACAGGCTGCTCTTCTCCAAATC 720
Qy	721 TGTAGCAAAACATAAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGT 780
Db	721 TGTAGCAAAACATAAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGT 780
Qy	781 CCCAGGAAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db	781 CCCAGGAAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Qy	841 ACAGATAGAGAAAGATGGAATGCGATAGGATGCAATATTTTGCAGACAGAAACGCTAG 900
Db	841 ACAGATAGAGAAAGATGGAATGCGATAGGATGCAATATTTTGCAGACAGAAACGCTAG 900
Qy	901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACAGTAAAGC 960
Db	901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACAGTAAAGC 960
Qy	961 AGCAATCGTACTGCTTAAGACACAAATAGCTGAAGCTTCAGAAAAAGTTCCCGACTCTCC 1020
Db	961 AGCAATCGTACTGCTTAAGACACAAATAGCTGAAGCTTCAGAAAAAGTTCCCGACTCTCC 1020
Qy	1021 AATTCTTCAAGAGCGGAAACAAATGGAATACAGGCTGAGAAAGATCTTAAATATCAAA 1080
Db	1021 AATTCTTCAAGAGCGGAAACAAATGGAATACAGGCTGAGAAAGATCTTAAATATCAAA 1080
Qy	1081 ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACA 1140
Db	1081 ACCTGCAGATGGTTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACA 1140
Qy	1141 AGGAGTACTATTGGTAGTATTGTTTCCATGCTCTTATAGATGATCTGAAATGAGAC 1200
Db	1141 AGGAGTACTATTGGTAGTATTGTTTCCATGCTCTTATAGATGATCTGAAATGAGAC 1200
Qy	1201 CGTTTCCATTTTGAATCTCGGTTTCGTAGATGATTACATGTTTCAATACGGAATATCC 1260
Db	1201 CGTTTCCATTTTGAATCTCGGTTTCGTAGATGATTACATGTTTCAATACGGAATATCC 1260
Qy	1261 TGAATCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGCAAGCCGCTGG 1320
Db	1261 TGAATCTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGCAAGCCGCTGG 1320
Qy	1321 AGATGACAGTCTGCTGCGAGCTGGCAGATGCTCAGAAAGCTTTAGAACGGCTCTAGG 1380
Db	1321 AGATGACAGTCTGCTGCGAGCTGGCAGATGCTCAGAAAGCTTTAGAACGGCTCTAGG 1380
Qy	1381 TAAAGCTGGGCAACAACAGGGCATACCTAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db	1381 TAAAGCTGGGCAACAACAGGGCATACCTAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Qy	1441 TGTGAGCGCAGGAGTCTCTCCCGCTGCGCAAGTTCTATAGGCTCATCTGTATAAACAGCT 1500
Db	1441 TGTGAGCGCAGGAGTCTCTCCCGCTGCGCAAGTTCTATAGGCTCATCTGTATAAACAGCT 1500
Qy	1501 TTACAACAGCTCAAAATCTACAGTCTGATTTAATAACACAGATATCAGCAGGTTATGA 1560
Db	1501 TTACAACAGCTCAAAATCTACAGTCTGATTTAATAACACAGATATCAGCAGGTTATGA 1560
Qy	1561 TGCTTACAAATCCATCAATGATGCTTATGTTAGGCAACGAAATGATCGGACTCTGATGT 1620
Db	1561 TGCTTACAAATCCATCAATGATGCTTATGTTAGGCAACGAAATGATCGGACTCTGATGT 1620
Qy	1621 GATAAACAATGTAAGTACCCCGCTCTCACACGATCGGTTCTTAGAGCACGAAACAGAGC 1680
Db	1621 GATAAACAATGTAAGTACCCCGCTCTCACACGATCGGTTCTTAGAGCACGAAACAGAGC 1680
Qy	1681 TCGAGGACCGAAGAAACACAGATCAAGCCCTCGTAGGTTGATTTCTGGCAATAGCAGAAC 1740
Db	1681 TCGAGGACCGAAGAAACACAGATCAAGCCCTCGTAGGTTGATTTCTGGCAATAGCAGAAC 1740

```
QY 1741 TCTTGAGATGCTCTAGTCAAGTTCGCGACTACATACTGTAATCGAGATCATCGATC 1800
Db 1741 TCTTGAGATGCTCTAGTCAAGTTCGCGACTACATACTGTAATCGAGATCATCGATC 1800
QY 1801 GAATCTCAAGCGAATAATGAGAGATCAGACAAAAGCTTACATCGGCGAGTCAAAAGCC 1860
Db 1801 GAATCTCAAGCGAATAATGAGAGATCAGACAAAAGCTTACATCGGCGAGTCAAAAGCC 1860
QY 1861 TCCACAGTTTGGCTATCTCTATGTCGAATCTTCTATGACTCTACACAGAAGTTCATAGC 1920
Db 1861 TCCACAGTTTGGCTATCTCTATGTCGAATCTTCTATGACTCTACACAGAAGTTCATAGC 1920
QY 1921 TAAATPAGAAAGTTTGTCTGCTGAAGATCTAGGACAGAGCTGTAATAAAGCACTTTC 1980
Db 1921 TAAATPAGAAAGTTTGTCTGCTGAAGATCTAGGACAGAGCTGTAATAAAGCACTTTC 1980
QY 1981 CTTTGAAGCGAATCTCTGTTTATTCAGCAGGTCTGCTCAATATCGGCTCTCTATATTC 2040
Db 1981 CTTTGAAGCGAATCTCTGTTTATTCAGCAGGTCTGCTCAATATCGGCTCTCTATATTC 2040
QY 2041 TGGTATCTCCATAACAACTAGTCTGTTTGGAGAGATTATTATGCTTTTGT 2100
Db 2041 TGGTATCTCCATAACAACTAGTCTGTTTGGAGAGATTATTATGCTTTTGT 2100
QY 2101 AAGGCTTTGTTGAGCCTTACCAACACTAGACGATCTTCAATAAATAAAGA 2156
Db 2101 AAGGCTTTGTTGAGCCTTACCAACACTAGACGATCTTCAATAAATAAAGA 2156

RESULT 2
AB001654/c
LOCUS amydia pneumoniae section 70 of 103 of the complete genome.
DEFINI 01654 AB001363
ACCESS 01654.1 GI:4377031
VERSIC
KEYWOR
SOURCE mydophila pneumoniae CWL029
ORGN

REFERENCE
1 (bases 1 to 10421)
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
AUTHORS Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R.W.,
Olinger, L., Grimwood, J., Davis, R.W. and Stephens, R.S.
TITLE Comparative Genomes of Chlamydia pneumoniae and C. trachomatis
JOURNAL Nat. Genet. 21 (4), 385-389 (1999)
MEDLINE 9920606
PUBMED 10192388
REFERENCE
2 (bases 1 to 10421)
Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Olinger, L.,
Grimwood, J., Davis, R.W. and Stephens, R.S.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1998) Program in Infectious Diseases, University
of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
FEATURES
Location/Qualifiers
1..10421
/organism="Chlamydia pneumoniae CWL029"
/mol_type="genomic DNA"
/strain="CWL029"
/db_xref="taxon:115713"
133..2751
/gene="CpN0727"
133..2751
/gene="CpN0727"
/codon_start=1
/transl_table=1
/product="CpN0727"
/protein_id="AADI8866.1"
/db_xref="GI:4377032"
/translacion="MSRKDNVSLARSIFNLSTGFSRITGIFREIMATYFGADPI
VAFWLGRVTFVRLKQIGLLEQAFIPHFELRAQSLDRAAFRRFRSLKIGSTI
IFTLIEAVLMVLQYVEEGYVETMLTLLPCGIFLMNVNAGALHLCNKFPGV
LAPVNNIWIFFVIAARHSDPRRIIGLSVALVGFWEFLITVPWKFLLEAKSP
POBHSVRALLAPLSGLITSSIFQLNLSICLARYVHEIGPLVSLYIOLPIH
LFGFGLVPLPAISRCVQREDHRLKMKFVLTLMSVMIINTAGLILALPGVRV
LYEHLPGQSAVAIVLRCYGASIIIPALAPLVLSULFYAQRQYAVPLFICGTALA
NIVLSLGRWLKDVSGISATISATVQVLIYFLWYSSKRLPMKLSWESIRRSIK
VWGTLMACMTLTLNLTITQTVYIFLNLPLPLAWPLUSSITAQIAFLSFCIFLAF
FGFAKLRLVEDLINLASFEYWRGQGLLQROHVQDTON"
8143..8319
/gene="CpN0731"
8143..8319
/gene="CpN0731"
8143..8319
/gene="CpN0731"

gene
CDS
```

```
1QAIDDAIKQAPALGGDGEIITIIQTLVPLVDKTTFTKAEFDLIYATQLPNTASLKL
YLTDQIAEYRKTKYQNSIONSETKRVNNRSMLETLQSLMFOQAQCFCVTTWIS
QANALNIAITNKYISAVLTSMEMYGGILCLSYMYERLADDEKAFIDKSNEYLIPIHI
VYGSWNGWIAKMAAYOELAEYSLGTAVTSDQDIKATIGRGNFKAHFFHNGID
QMGYPANETVFGNLTITANGAIQPDGGFIREAMTNVGTVEADYVSNARQIRILNEFTA
ATPHVQLQIAELQKQADDLPKASFTENRKFVAAWITSESLGALISMLNSQ
LPQEAFLIPLIEINFNLNANLNSLQITNEFTSTVYSSLSYLVQSTQGNLF
AGDYVETLAAAREYIYRDTARCKOAILNVLNKLKINSIPGATSKQKOBMLNATF
YQYSLSTNOLTVLESLAGLKWTLQTTNNGKYDKSVKTESFDDWIPTLAALESF
LTSFPNLSATGGGLGPLEFTVQSQOQTYTSQOTQOQLNLQNTTIQEWLIVSTSMQ
VNLGLSLQAGAIYSN"
complement (3078)..5033
/gene="CpN0728"
complement (3078)..5033
/gene="CpN0728"
/codon_start=1
/transl_table=1
/product="CHLPN 76 kDa Homolog_1 (CT622)"
/protein_id="AADI8867.1"
/db_xref="GI:4377033"
/translacion="WNPFGPIDEETERTPPADLSAOGLEASRANKSAEQRLAGAE
AKPKESKDSVERWSILRSVNALMSLADKLGIASSNSSSTSRSDVDSITATPTP
PPFPDDYKTAQATYDTIFTSTSLADIQAALVSLQDRAVTKIQTAADEETATAAEW
ETKNADAVKVQAITEKAYASDNQALDLSKLITFDLLQAALLQSVANNKAABL
KEMQNPVPGKTPAIQSLVDQDATATQIEKQGNARIDAYFAGNAGAVNAKSN
NSINIDSAKAIATAKTQIAEAKQKPPDSPILOAEQMTIOAKDLKNIKPQSDV
PNPQTVGSKQGGSSIRVSLMLDDAENETASILMSGFQMTLHMENTENPDSQAA
OQEELAAQARAAGDSDAALADAKALEAALGKAGQOQOGLNALQIASAAYVSA
GVPPAASSIGSSVKLYKTSKSDYKQISAGYDAYKSDINAYGARNDATRDVI
NNVSTPALTRSPRPAETKPEKTDQALARVISGNSRTLGVSQVSAQSVMQIIC
SNPQANNEIRKLTSAVTKPPQFYQVYVQLSNDSTQKFIARLESFLEAGSRTAAEIK
ALSTPNSLFTIQVVLNIGSLYGLQ"
complement (5145)..6413
/gene="CpN0729"
complement (5145)..6413
/gene="CpN0729"
/codon_start=1
/transl_table=1
/product="CHLPN 76 kDa Homolog_2 (CT623)"
/protein_id="AADI8868.1"
/db_xref="GI:4377034"
/translacion="WKQGVYQWLASVLLALTISGYAELPLSEQKVKSHYTTTLDVK
DYLKRGFVETRKQDGLRIAGDVRARWLIFREDIKNPSDKDKYNPLFVNRYSFYL
YIDRAERNWLSKMNMTAIGAGNTAGVDINRAFLYKPKETDPTDFMEIGRS
GLGLDFSEVQFQSGNDGLHTYWTRELSKDPYQVIVHGGFPVVMNTKKHVAWVEGI
LNRLPKQFVKCSVDWNTVPSTSTTEKATAMKYKVCVQWOLVSKHSQVPIWG
OKPLYLNGALMPLAKATKTLNGKENLAWFLGTLGLKRLKQKQWATVREYVEA
LSVEIDVSGIGRGNLLKPFQAQALAAANYDPEANGFTNYKGFSLYNYGTTDSLSFR
AIGAYSKPANDKLGSDFTFRKFDGILISAF"
complement (6516)..8159
/gene="mvln"
/notes="synonym: CpN0730"
complement (6516)..8159
/gene="mvln"
/codon_start=1
/transl_table=1
/product="Integral Membrane Protein"
/protein_id="AADI8869.1"
/db_xref="GI:4377035"
/translacion="MSRKDNVSLARSIFNLSTGFSRITGIFREIMATYFGADPI
VAFWLGRVTFVRLKQIGLLEQAFIPHFELRAQSLDRAAFRRFRSLKIGSTI
IFTLIEAVLMVLQYVEEGYVETMLTLLPCGIFLMNVNAGALHLCNKFPGV
LAPVNNIWIFFVIAARHSDPRRIIGLSVALVGFWEFLITVPWKFLLEAKSP
POBHSVRALLAPLSGLITSSIFQLNLSICLARYVHEIGPLVSLYIOLPIH
LFGFGLVPLPAISRCVQREDHRLKMKFVLTLMSVMIINTAGLILALPGVRV
LYEHLPGQSAVAIVLRCYGASIIIPALAPLVLSULFYAQRQYAVPLFICGTALA
NIVLSLGRWLKDVSGISATISATVQVLIYFLWYSSKRLPMKLSWESIRRSIK
VWGTLMACMTLTLNLTITQTVYIFLNLPLPLAWPLUSSITAQIAFLSFCIFLAF
FGFAKLRLVEDLINLASFEYWRGQGLLQROHVQDTON"
8143..8319
/gene="CpN0731"
8143..8319
/gene="CpN0731"
8143..8319
/gene="CpN0731"
```

```
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAD18970.1"
/db_xref="GI:4377036"
/translation="MSFLLIKPFSPPLKQDYLFDISPTSSEITIGGSYFKNKASL
QSTLRRLRSISIIIS"
gene
8654..9535
/genes="nfo"
/feature="synonym: CPn0732"
8654..9535
/genes="nfo"
/feature="nfo"
/codon_start=1
/transl_table=11
/product="Endonuclease IV"
/protein_id="AAD18971.1"
/db_xref="GI:4377037"
/translation="MKVLPPPSIPLLGAHTSTAGGLKNAIYGRDIGASTVQIFTAHQ
ROWRRALKBEVIEDFAALKETDLSYIMSHAGYLINPGAPDFVILEKRIYIYEIL
DCITLISFVNPHGAALKSKEDCMNKIVSFQSAPLFDSPFLVLELLETTAGGT
LIGSNFELGYLVQNLKQIPIGVCVDTCHIIPAAGYDITSPQWBDVLNERPEYVGLS
YLRAHLNDSMPPLGANKDRHAPLGBGYIGKESFKFLMTDERTKIPKYLETPGPEP
WKEIGELLKFSKNRDS"
gene
complement(9654..10283)
/genes="rs4"
/feature="synonym: CPn0733"
complement(9654..10283)
/genes="rs4"
/codon_start=1
/transl_table=11
/product="S4 Ribosomal Protein"
/protein_id="AAD18972.1"
/db_xref="GI:4377038"
/translation="WARYCGPKNRVARRFGANI FGRSRNPLLKHPBPQOHGMORKK
SDYGLQLEKOKLKACYGMIMEKOLVAKPEVHKQGNVAQWELERFPCRLDNNYRM
GFAPTLFAAQQLVAHGHILVNGRRVDRSFLRPGMQISLKESKRLQSVKDALRESKD
ESSLSYISLSDXGFGKELLVSPEDQIEAQLPILNISVVCFLSHRT"
ORIGIN
Query Match 100.0%; Score 2156; DB 1; Length 10421;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTGAGAGCTTTTATTTT 60
DB 5133 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTGAGAGCTTTTATTTT 5074
QY 61 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTGAGAGCTTTTATTTT 120
DB 5073 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTGAGAGCTTTTATTTT 5014
QY 121 AGGTCTTATAGCAAGAACGACACCTCCGCGAGATCTTCTGCTCAAGGATTGGA 180
DB 5013 AGGTCTTATAGCAAGAACGACACCTCCGCGAGATCTTCTGCTCAAGGATTGGA 4954
QY 181 GCGAGTGCAGCAATATAGATGCGGAAGCTCAAGAAATAGCAGTGCGGAAGCTTAAGCC 240
DB 4953 GCGAGTGCAGCAATATAGATGCGGAAGCTCAAGAAATAGCAGTGCGGAAGCTTAAGCC 4894
QY 241 TAAAGAACTTAAGACCGATTCTGTAGAGCGATGAGAGCATCTTGCGTCTGAGTGAATGC 300
DB 4893 TAAAGAACTTAAGACCGATTCTGTAGAGCGATGAGAGCATCTTGCGTCTGAGTGAATGC 4834
QY 301 TCTCATAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAG 360
DB 4833 TCTCATAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAG 4774
QY 361 CAGATCTCAGACGTGGACTCAACAGACGACGACCGCACTTACGCTCTCTCCACCAAGTT 420
DB 4773 CAGATCTCAGACGTGGACTCAACAGACGACGACCGCACTTACGCTCTCTCCACCAAGTT 4714
QY 421 TGATGATTATTAAGACTCAAGCGCAACAGCTTACGATACTATCTTACTCTCAACATCACT 480
DB 480
```

1561 TGCTTCAAAATCCATCAATGATGCTATGTTAGGACGAAATGATCGAGCTGCTGATGT 1620
Db TGCTTCAAAATCCATCAATGATGCTATGTTAGGACGAAATGATCGAGCTGCTGATGT 3514
1621 GATAAACAATGTAAGTACCCCGCTCTCACAGATCGCTTCTAGACGAGAACGAAGC 1680
Db GATAAACAATGTAAGTACCCCGCTCTCACAGATCGCTTCTAGACGAGAACGAAGC 3454
1681 TCAGGACACGAGAAACAGATCAAGCCCTCGTAGGCTGATTTCTGGCAATAGCAGAAC 1740
Db TCAGGACACGAGAAACAGATCAAGCCCTCGTAGGCTGATTTCTGGCAATAGCAGAAC 3394
1741 TCTTGGAGATGCTATAGTCAAGCTTTCGGCACTACCAATCTGTAATCGAGATCATCCAGTC 1800
Db TCTTGGAGATGCTATAGTCAAGCTTTCGGCACTACCAATCTGTAATCGAGATCATCCAGTC 3334
1801 GAATCTCAAGCAATATAGGAGATTCAGCAAAAGCTTACATCGGCAGTGCAAAAGCC 1860
Db GAATCTCAAGCAATATAGGAGATTCAGCAAAAGCTTACATCGGCAGTGCAAAAGCC 3274
1861 TCACAGATTTGGCTATCTTATGCTCAACTTCTTAATGACTCTACACAGAGTTCATAGC 1920
Db TCACAGATTTGGCTATCTTATGCTCAACTTCTTAATGACTCTACACAGAGTTCATAGC 3214
1921 TAAATAGAAAGTTCTTGTGCTGAAGATCTAGCAGCAGCTGAAATAAAGCACTTTC 1980
Db TAAATAGAAAGTTCTTGTGCTGAAGATCTAGCAGCAGCTGAAATAAAGCACTTTC 3154
1981 CTTTGAAGCAACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db CTTTGAAGCAACTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3094
2041 TGGTATCTCCATAACCAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db TGGTATCTCCATAACCAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3034
2101 AAGGCTTTGTTGAGGCTTACCAACACACTAGCAGCTTCTCAATAAATAAAGA 2156
Db AAGGCTTTGTTGAGGCTTACCAACACACTAGCAGCTTCTCAATAAATAAAGA 2978

RESULT 3
LOCUS AE002165 12548 bp DNA linear BCT 30-MAY-2000
DEFINITION Chlamydia pneumoniae AR39, section 2 of 94 of the complete genome.
ACCESSION AE002165
VERSION AE002165.1 GI:7188948
KEYWORDS
SOURCE Chlamydia pneumoniae AR39
ORGANISM Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwinn, M., Nelson, W., deBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
TITLE Genome sequences of Chlamydia trachomatis Mogen and Chlamydia pneumoniae AR39
JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)
MEDLINE 20150255
PUBMED 10684935
REFERENCE 2 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwinn, M., Nelson, W., deBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
MEDLINE Medical Center Dr, Rockville, MD 20850, USA

FEATURES
Source
1. 12548
/organism="Chlamydia pneumoniae AR39"
/mol_type="genomic DNA"
/strain="AR39"
/db_xref="taxon:115711"
/note="synonym: Chlamydia pneumoniae AR39"
297. 965
/gene="CP0011"
297. 965
/gene="CP0011"
/note="similar to SP:P31218 GB:X71492 PID:296947 GB:U00096
PID:1736770; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="uridine kinase"
/protein_id="AAF37907.1"
/db_xref="GI:7188950"
/translation="MLMMLMIIGITGGGAGKTTLTQNIKEIFGEDSVICODNVYK
DRSHVTPPERANLIWDHDAFDNDLLISDIKELKNEIIVQAPVDFVLGNSKTRLET
LYPSXVILVEGLVFEQELDMDIRIFVDTDADERILRMVRDVQGGSDVDCIMS
RYLSWVKEHKEFIEPTKRYADIIVHGNRYQNVVITNLSQIKHLENALSDSETIYM
VNSK"
complement (1047. .2021)
/gene="CP0012"
complement (1047. .2021)
/gene="CP0012"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37908.1"
/db_xref="GI:7188951"
/translation="MEKYLALAYITRVNPHIEALHKKFLLEDLVSCRIYISEQ
GNGQFSGYEPAELVQWMLKERNFNSKFIHIIKENIFRITVKRKLALGCE
VLSQAKHISPEQWAEKLEQNRCLILVNNYEWKIGHFNATLPDIQTREFPEYKA
EKLAGCDEPTTVMWCTGIGRCELYSPLEKFEVYOLDGVIAYGOQVGTGW
LGLFVFDRLAIPIDASDPVATACCHCQTPSDAYNCANTCNALFCCDBCIH
HQGCCSECSQSFVRVAFDSRGNKPPRRHLCEISENSEASCLLI"
2238. .2867
/gene="CP0013"
2238. .2867
/gene="CP0013"
/note="similar to SP:P81288; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S4"
/protein_id="AAF37909.1"
/db_xref="GI:7188952"
/translation="MAYCYGPKRVARRFGANIFGRSENPLKKPHPPGQHGMMKKK
SDYGLQREKQIKACYGMINEKQIVAKFVHKQNVAGQMFLEFECRLDNNVYRM
GKATITAAQQLVAHGHILVNGRRVDRSPFLRPMQISLKEKSKLQSVDALESKD
ESSLPYSILDKTGKGLVSPEDQTEAQLPLINISVCEFLSHRT"
complement (2986. .3867)
/gene="CP0014"
complement (2986. .3867)
/gene="CP0014"
/note="similar to SP:P12638 GB:M22591 PID:146954
PID:405898 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="endonuclease IV"
/protein_id="AAF37910.1"
/db_xref="GI:7188953"
/translation="MKVLPPPIPLLGAHTSTAGLKNATVEGRDIGASTVOITFANO
ROMORALKSEVIEDFAALKETLSVIMSHAGVLIINGAPDPVILEKSLIGIYQEL
DCITLGSFNVNHFCAALKSKEDCMNKIVSSFSQAPLFDSSPLVLLVLTAGQEL
LYGFNFBELGLVQNLKNQIPIGVCDTCHIFAAGYDITSPQGWEDVNBFEYVGLS
LIFRAHLNDSMFLGANKRHRAPLIGEGYIGKESFPLMTDERTKI PKYLETPGPPN
WQKEIGELLKFSKNRDS"

QY 661 TCTTGACCTTTTGGTAAACTGACTTCTTCTGACCTCTTACAGGCTGCTCTTCTCAATC 720
Db 8048 TCTTGACCTTTTGGTAAACTGACTTCTTCTGACCTCTTACAGGCTGCTCTTCTCAATC 8107
QY 721 TGTAGCAACAATATCAAGAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db 8108 TGTAGCAACAATATCAAGAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 8167
QY 781 CCCAGGGAACCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 8168 CCCAGGGAACCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 8227
QY 841 ACAGATAGAAAGATGGAATGCGAATAGGATGATATTTTGAGGACAGAACGCTAG 900
Db 8228 ACAGATAGAAAGATGGAATGCGAATAGGATGATATTTTGAGGACAGAACGCTAG 8287
QY 901 TGGAGCTGTAGAAATGCTAAATCTTAATACAGATATAGCAACATAGATTAGCTAAAGC 960
Db 8288 TGGAGCTGTAGAAATGCTAAATCTTAATACAGATATAGCAACATAGATTAGCTAAAGC 8347
QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 1020
Db 8348 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC 8407
QY 1021 AATTTCTCAAGAGCGGAACAATGGTAAATACAGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 8408 AATTTCTCAAGAGCGGAACAATGGTAAATACAGCTGAGAAAGATCTTAAAAATATCAA 8467
QY 1081 ACCTGCAGATGCTGTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAGCAACA 1140
Db 8468 ACCTGCAGATGCTGTGATGTTCCAAATCCAGAACTACAGTTGAGGCTCCAGCAACA 8527
QY 1141 AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 1200
Db 8528 AGGAAGTAGTATTGGTAGTATTGCTGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 8587
QY 1201 CGCTTCCATTTTGTGCTGCTGGTTTTCGTCAGANGATTACATGTTTCAATACGGAATTC 1260
Db 8588 CGCTTCCATTTTGTGCTGCTGGTTTTCGTCAGANGATTACATGTTTCAATACGGAATTC 8647
QY 1261 TGATTTCAAGCTGCCCAACAGAGCTCGCAGCACAAAGCTAGCAGCAAGCGCGTGG 1320
Db 8648 TGATTTCAAGCTGCCCAACAGAGCTCGCAGCACAAAGCTAGCAGCAAGCGCGTGG 8707
QY 1321 AGATGACAGTGTCTGTCAGCGCTGCGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
Db 8708 AGATGACAGTGTCTGTCAGCGCTGCGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 8767
QY 1381 TAAAGCTGGGCAACAACAGGGGATCTCAATGCTTTAGACAGATCGCTTCTGCTGTGT 1440
Db 8768 TAAAGCTGGGCAACAACAGGGGATCTCAATGCTTTAGACAGATCGCTTCTGCTGTGT 8827
QY 1441 TGTAGCGCAGAGTTCCTCCCGCTGCGCAGAGTTCTATAGGTCATCTGTAAACAGCT 1500
Db 8828 TGTAGCGCAGAGTTCCTCCCGCTGCGCAGAGTTCTATAGGTCATCTGTAAACAGCT 8887
QY 1501 TTAAGAAGCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGA 1560
Db 8888 TTAAGAAGCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGA 8947
QY 1561 TGCCTTCAAAATCCATCAATGATGCTTATGTTAGGACAGAAATGATGCACTCGTAGTGT 1620
Db 8948 TGCCTTCAAAATCCATCAATGATGCTTATGTTAGGACAGAAATGATGCACTCGTAGTGT 9007
QY 1621 GATAAACAATGTAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAACAAAGAGC 1680
Db 9008 GATAAACAATGTAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAACAAAGAGC 9067
QY 1681 TCGAGGACCAAGAAAAACAGATCAAGCCCTCGTAGGTTGATTTCTGGCAATAGCAGAAC 1740
Db 9068 TCGAGGACCAAGAAAAACAGATCAAGCCCTCGTAGGTTGATTTCTGGCAATAGCAGAAC 9127

QY 1741 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGAGATCATCCAGTC 1800
Db 9128 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGAGATCATCCAGTC 9187
QY 1801 GAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATGGCAGTGACAAAAGC 1860
Db 9188 GAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATGGCAGTGACAAAAGC 9247
QY 1861 TCCACAGTTTGGCTATCCTTATGTGCAACTTCTAATGACTCTACACAGAAGTTTCATAGC 1920
Db 9248 TCCACAGTTTGGCTATCCTTATGTGCAACTTCTAATGACTCTACACAGAAGTTTCATAGC 9307
QY 1921 TAAATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAAATATAAAGACACTTTC 1980
Db 9308 TAAATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAAATATAAAGACACTTTC 9367
QY 1981 CTTTGAAACGAACCTCTTGTGTTTATTCAGCAGGCTGCTCAATATCGGCTCTCTATATTC 2040
Db 9368 CTTTGAAACGAACCTCTTGTGTTTATTCAGCAGGCTGCTCAATATCGGCTCTCTATATTC 9427
QY 2041 TGGTTATCTCCAATAACAACACCTAAAGTTTCGTTTGGAGAGATTATATGCTTTCGT 2100
Db 9428 TGGTTATCTCCAATAACAACACCTAAAGTTTCGTTTGGAGAGATTATATGCTTTCGT 9487
QY 2101 AAGCCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 2156
Db 9488 AAGCCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 9543

RESULT 4
AE017159/c
LOCUS
DEFINITION
Chlamydomophila pneumoniae TW-183, section 3 of 4 of the complete genome.
ACCESSION
AE017159 AE009440
VERSION
AE017159.1 GI:33236383
KEYWORDS
SOURCE
ORGANISM
Chlamydomophila pneumoniae TW-183
Chlamydomophila pneumoniae TW-183
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia. phila.
REFERENCE
1 (bases 1 to 300066)
Geng, M.M., Schumacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T., Essig, A., Marre, R. and Melchers, K.
The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 300066)
Geng, M.M., Schumacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T. and Melchers, K.
Direct Submission
JOURNAL
Submitted (28-MAY-2002) RDR/IT, RDR/FG, RDR/BT, ALTANA Pharma, Byk-Gulden-Str. 2, Konstanz 78467, Germany
FEATURES
source
1 300066
/organism="Chlamydomophila pneumoniae TW-183"
/mol_type="genomic DNA"
/strain="TW-183"
/db_xref="taxon:182082"
/note="type strain"
biovar: "TWAR"
126..797
/locus_tag="CpB0543"
126..797
/locus_tag="CpB0543"
/locus_tag="CpB0543"
/note="similar to hypothetical protein CT433 - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98472.1"
/db_xref="GI:33236384"


```

/translation="MTLYLGLNOKTARKQAOHYLPILTLFPYAKSTPONKRALQELPO
ATHVILTSSSTHLFLSRMTSLSKATLTKTKVLGICESTKSKRLSLFLQGVKYVATQ
ETAEGIFLLOALPSSARILYPHSSILARPIWIEFLYNRTFFSYHYITVKKRLKXNI
LSYKXLIITSPSTVRAFAKIPFRFFPEKTYWCQGRMTLQEFQKFSQKQVSLLETIGK
SRTSP"
gene      862..1194
/locus_tag="CpB0544"
CDS      862..1194
/locus_tag="CpB0544"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98473.1"
/db_xref="GI:33236385"
/translation="MASSATPGDGTAPSLEPLATEPRYNFKLAFVITIALVWIAL
IATIIAIGLHCIFLCSFILTALPIFLIFSRICSHYARNVYIALDVDPHSLQDMRS
HSIFSDR"
gene      1242..2321
/locus_tag="CpB0545"
CDS      1242..2321
/locus_tag="CpB0545"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98474.1"
/db_xref="GI:33236386"
/translation="MSGPQRTSSQSVLSYVPRDKEIAPKKQFTIAKISTLAILASL
ALGALVAGISLITVILGNVFLALLITLALFVSFVTLVHOMSKVSNWQKVLQGNFK
PLGKQERQNVDCYNEMQFYANHLNPKFKVAILQIDASQFPFTLTGLRVLKRNQST
GLIFNPQTNLIDNTATNLSTILYSTLKDKSVNCTKOREGDPKAGEDPPFPSTVRV
KLIPEALQDTNMLNSAAEKSLILPTFLGHVCGPKSELPNQOBYRQALLAYENCL
KAAITSHAIVALEPFTVSYVEVPEPPEILPKEGTFYWDNQTQCFKRALDLAIONTAR
YQPSLLVLPQPFNTIESQSRSE"
gene      2661..3425
/locus_tag="CpB0546"
CDS      2661..3425
/locus_tag="CpB0546"
/translation="similar to hypothetical protein CT398 - Chlamydia
trachomatis"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAP98475.1"
/db_xref="GI:33236387"
/translation="MHDALLSILAIQELDIKMLRMVKKHOKELAKVQSLKSDIR
KYQKELEMENLKTQIRDCENRIQISQINKLENOAAVKMDFNALTOEMTTANK
ERRLEHOLDLMDKQAGEDLIVSLKESLASTENS SVIEKEIFESIKKINEEGKAL
LEORTELKATNPILLSIYERLLNNKKRVVPIENRVCSGCHIVLTPOHENLVKKKD
RLIFCEHSRLIYWQESQVNAQENSTAKRRRRRAV"
gene      3846..4835
/locus_tag="CpB0547"
CDS      3846..4835
/locus_tag="CpB0547"
/translation="similar to kpsF protein - Escherichia coli"
/codon_start=1
/transl_table=11
/product="kpsF"
/protein_id="AAP98476.1"
/db_xref="GI:33236388"
/translation="MPSPMISTVQDILGKQKAEVDFFOAQFQKMAQLAEKILGH
SGWFFGVGSGCKVARKVATLQSLSERALFFSPVDLLHGLDGLVSPGDIVCLFSSK
GETQLLTLPKLSRAILVAITSMYPYSLNLAASDLVILPSVAELDPFNLIPTNST
TCQMIFGFLAMLLPFLKSGVSLSTYGNKHPSGQGMKANGKVDKPMPEKTVPECHLG
DKVPSFLVFSYAGCGVCIVDPQPLMGIITDGLRSLRSYSGEVLISLEKLVMTA
NPRCITESDIALQLMESSPVALPVLONEENRHVTLGHMHTLAKAGLL"
gene      4862..6043
/locus_tag="CpB0548"
CDS      4862..6043
/locus_tag="CpB0548"
/translation="similar to dihydrolipoamide S-acetyltransferase -
Bacillus stearothermophilus"

```

```

/codon_start=1
/transl_table=11
/product="dihydrolipoamide S-acetyltransferase"
/protein_id="AAP98477.1"
/db_xref="GI:33236389"
/translation="WIFPRFPKIGETSSGGSIVFWLKNLGDHVARDPLEIVSTDKI
ATELPSRAGRLVRCVNEGDEVAGDVLGIELEIEISEADDESTCPLTSCETKSEA
GSSSSVWSPAVLSLAQREGIGLDNLQIAGTKGGRVTRQDLQEIYSESQQVQIPE
IFQGEVNRIPMSPLRRRAIASLSKSSDEVPASLVVDVDTLMLHISGERQFLDTH
GVKLTITSIVQCLAQTLRQFLPLNGSLDGTIIVMKKSVNVGVANLNKGEVVPVIV
NQCDRLVSIKALADLSSRLRNKLDPSEVQDGSVTVTNFGMTGALIGPILIRPEV
AALLGIGTIQKRWVVRDDDSLAVRKMYVYVLTIFDHRVLGDIQSEPLTSLKRLSVTM
G"
gene      6057..7301
/locus_tag="CpB0549"
CDS      6057..7301
/locus_tag="CpB0549"
/translation="similar to proton/sodium glutamate symport
transporter family"
/codon_start=1
/transl_table=11
/product="proton/sodium glutamate symport transporter"
/protein_id="AAP98478.1"
/db_xref="GI:33236390"
/translation="MKLWKIFIGLVGVTLGLVLEDKAIFPKIGDIFLMLSMVY
PLVCSVMVLGASISDMKKLGRIGISVGLXILGTTALAIIVIGLCFAMIFSPNGCDFA
QAQMSDASVTVIDSNKTAAYFLSIIAQVFPSPNVRSPAEGNLIQIIIFAIFIGIALRL
SGERPRFERIDFGESEIMLRVMNIMSFAPYGVGASMAWISGNHGLVGLWLGKFII
AYILACIFHATFVGFQKMSFSLSSMDMAISCAVSTASSATLPTVMRCVS
KNLGSABVSGFVLPLGATVMNMTAIFQGMAAVFIACAYNCPISLSLILVVTATF
SALVSGAVPGSGMITLGSVLASVGLPIQGIALLAGIDRLRIVGTFMNIGDAVVATY
VASEGELSPYESIKQSVET"
gene      7301..8398
/locus_tag="CpB0550"
CDS      7301..8398
/locus_tag="CpB0550"
/translation="similar to ycaH protein - Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="ycaH"
/protein_id="AAP98479.1"
/db_xref="GI:33236391"
/translation="MKRFPSTLFLFVRVTIALLSGLGWMGSLLSKVAFPLVA
CWNFSTPTFRARSTVISGNIIVVGAGKTPPTVLAEALRLRGYSGVLSRGYKQ
SSRQKLTIVDSKVHSASYGDEPLMAEKLPEGSVWVKDRRIISAAAEKFGILL
DDGIQYKLHKDVEIAVNGQDPLGGRAFPFGRLRDFPLRLKTVDALIVNGGKKEAG
TVKRVSNAPQIFVKPTIASVWTHNGERIPKEALRELVRGVFCGLGPGQFLNMLE
EGIHILGKYLLOHAALITKKELNYFCOAMARQOGLLCTEKOSVKLPRLSGEVSLIP
IAKVMRLSVNQDDTLNLNKHQHN"
gene      8596..9393
/locus_tag="CpB0551"
CDS      8596..9393
/locus_tag="CpB0551"
/translation="similar to trna methylase spoU - Escherichia coli"
/codon_start=1
/transl_table=11
/product="trna methylase"
/protein_id="AAP98480.1"
/db_xref="GI:33236392"
/translation="MDCIGKHNPLVEALALKSRCKSKSWFLVEGAREIQKALRTGY
LCOHFVCSHLSSEKKEFLYELKRNSTKILYCLDSTLAQLSPKEHDSFVAFQKRW
NKEDFLTKRNAQPFYLIIEQVEKPNVGAILEIADGAGVDGVIDLCNPIVDLYNPV
RSSGLVFSPLISISREGEKLFQEGMTVFTSPRAETMYFSKNVLTALVFGSE
KDGLEDWFESEFSEIAPLMIGESDSINLATSVAAYEVVVRQW"
gene      9381..10208
/locus_tag="CpB0552"
CDS      9381..10208
/locus_tag="CpB0552"
/translation="similar to probable SAM dependent methyltransferase"

```



```
- Chlamydia trachomatis"
/codon_start=1
/transl_table=11
/product="putative SAM dependent methyltransferase"
/protein_id="AAP98481.1"
/db_xref="GI:33236393"
/transl_table="MDKLLDSGDNKLECFGPVTLIRPSSIAVWPKSRPELWSAQL
QYVEGREGKAKNFKRLPEWEVAFSDVRCLLKFTPEHGLGVPEHMGFWPALKQALIE
KHKEQVNLNFAVTAGSIEFAAKCQARVTHVDASQAAVRAQNRVKNAPERRIFWV
IEDVISLKEIRNKKYQVILLDPSPYGRGPGEVFKIDKLDPLPPLSLSKLLADA
SYELLTSHTEPFLRATARSVPVTLVSEANSGESFGEVGVGALPSSSFQWIA"
complement(10211..10813)
gene
Query Match 100.0%; Score 2156; DB 1; Length 300066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATAAATCTTTAAACAGGCTGCATTAATTTATTTTGGATTTTATTTTATTTTATTTT 60
Db ATAAATCTTTAAACAGGCTGCATTAATTTATTTAGTAGAGCTTTTATTTTATTTT 214642
QY 61 ATATTAACCTAAAGATTTTATTTATTTTGGATTTTATTTTATTTTATTTTATTTT 120
Db ATATTAACCTAAAGATTTTATTTATTTTGGATTTTATTTTATTTTATTTTATTTT 214582
QY 121 AGTCTCTATAGAGAAACAGAACGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 180
Db AGTCTCTATAGAGAAACAGAACGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 214522
QY 181 GCGAGTGCAGCAAAATAGAGTCCGAGCTCAAGNATAGCAGGTGCGGAAGCTAAGCC 240
Db GCGAGTGCAGCAAAATAGAGTCCGAGCTCAAGNATAGCAGGTGCGGAAGCTAAGCC 214462
QY 241 TAAAGATCTAAGACCGGATCTCTAGAGCGATGGAGCATCTTCGGTCTCGAGTGAATGC 300
Db TAAAGATCTAAGACCGGATCTCTAGAGCGATGGAGCATCTTCGGTCTCGAGTGAATGC 214402
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATGTTCTTAGTAAACAGCTCGTCTTCTACTAG 360
Db TCTCATGAGTCTGGCAGATAAGCTGGGTATGTTCTTAGTAAACAGCTCGTCTTCTACTAG 214342
QY 361 CAGATCTGACAGCTGGATCTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTT 420
Db CAGATCTGACAGCTGGATCTCAACGACAGCGACCGACCTACGCTCTCCACCCACGTT 214282
QY 421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 480
Db TGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTTTTACCTCAACATCACT 214222
QY 481 AGCTGACATACAGGCTGTTTGGTAGCCTCCAGGATGCTGTCACTAATATAAGGATAC 540
Db AGCTGACATACAGGCTGTTTGGTAGCCTCCAGGATGCTGTCACTAATATAAGGATAC 214221
QY 541 AGCGGCTACTGATGAGAAACCGCAATCGCTGGAGTGGGAAACTAAGATGCCGATGC 600
Db AGCGGCTACTGATGAGAAACCGCAATCGCTGGAGTGGGAAACTAAGATGCCGATGC 214162
QY 601 AGTTAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGTTTCGGAATAACCAAGCGAT 660
Db AGTTAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGTTTCGGAATAACCAAGCGAT 214102
QY 214101 AGTTAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGTTTCGGAATAACCAAGCGAT 214042
QY 661 TCTTGACTCTTTAGGTAACTGACTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 720
Db TCTTGACTCTTTAGGTAACTGACTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAATC 213982
QY 721 TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGT 780
Db TGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGT 213922
QY 781 CCCAGGAAAACCGCTGCAATTTGCTCAATCTTTTAGTTGATCAGACAGATGCTTACAGCGAC 840
Db CCCAGGAAAACCGCTGCAATTTGCTCAATCTTTTAGTTGATCAGACAGATGCTTACAGCGAC 213862
```

```
QY 841 ACAGATAGAGAAAGATGGAATCGATTAGGATGCAATATTTTTCAGGACAGAACCGTAG 900
Db 213861 ACAGATAGAGAAAGATGGAATCGATTAGGATGCAATATTTTTCAGGACAGAACCGTAG 213802
QY 901 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAAGTAAAGC 960
Db 213801 TGGAGCTGTAGAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTCAAGTAAAGC 213742
QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACCTCTCC 1020
Db 213741 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACCTCTCC 213682
QY 1021 AATCTTCAAGAACGCGAACAAATGCTAATACAGGCTGAGAAAAGTCTTAAATAATCAA 1080
Db 213681 AATCTTCAAGAACGCGAACAAATGCTAATACAGGCTGAGAAAAGTCTTAAATAATCAA 213622
QY 1081 ACCTGAGATGGTTCGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 1140
Db 213621 ACCTGAGATGGTTCGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 213562
QY 1141 AGGAAGTACTATTTGGTAGTATTTGCTGCTTTCCATGCTGTTAGATGATGCTGAAATGAGAC 1200
Db 213561 AGGAAGTACTATTTGGTAGTATTTGCTGCTTTCCATGCTGTTAGATGATGCTGAAATGAGAC 213502
QY 1201 CGCTTCCATTTTATGCTGCTGGGTTTCGTGATGATGATTTCAATGCTTCAATACGGAATCC 1260
Db 213501 CGCTTCCATTTTATGCTGCTGGGTTTCGTGATGATGATTTCAATGCTTCAATACGGAATCC 213442
QY 1261 TGATTTCTAAGCTGCGCAACAGAGCTCGCAGCACAAAGCTAGAGCAGCAAGACCCGCTGG 1320
Db 213441 TGATTTCTAAGCTGCGCAACAGAGCTCGCAGCACAAAGCTAGAGCAGCAAGACCCGCTGG 213382
QY 1321 AGATGACAGTGTGCTGCGACGCTGCGACATGCTCAGAAAGCTTTAGAAAGGCTCTAGG 1380
Db 213381 AGATGACAGTGTGCTGCGACGCTGCGACATGCTCAGAAAGCTTTAGAAAGGCTCTAGG 213322
QY 1381 TAAAGCTGGCGCAACACAGGCGATCTCAATGCTTTAGACACAGATCGCTTCTGCTGCTGT 1440
Db 213321 TAAAGCTGGCGCAACACAGGCGATCTCAATGCTTTAGACACAGATCGCTTCTGCTGCTGT 213262
QY 1441 TGTGAGCGCAGAGTTCCTCCGCTGCGACAAAGTTCTATAGGCTCATCTGTAAAAAGCT 1500
Db 213261 TGTGAGCGCAGAGTTCCTCCGCTGCGACAAAGTTCTATAGGCTCATCTGTAAAAAGCT 213202
QY 1501 TTACAGAGCTCAAAATCTACAGGTTCTGATTTAAACACAGATATCAGCAGGTTATGA 1560
Db 213201 TTACAGAGCTCAAAATCTACAGGTTCTGATTTAAACACAGATATCAGCAGGTTATGA 213142
QY 1561 TGCCTTACAAATCCATCAATGATGCTGATGGGCAAGAAATGATGCGACTCGTGATGT 1620
Db 213141 TGCCTTACAAATCCATCAATGATGCTGATGGGCAAGAAATGATGCGACTCGTGATGT 213082
QY 1621 GATAAACAAATGTAAGTACCCCGCTCTCAGAGTCCGTTCTAGACACGAAAGAGC 1680
Db 213081 GATAAACAAATGTAAGTACCCCGCTCTCAGAGTCCGTTCTAGACACGAAAGAGC 213022
QY 1681 TCGAGGACACAGAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGCGCATAGCAGAAC 1740
Db 213021 TCGAGGACACAGAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGCGCATAGCAGAAC 212962
QY 1741 TCTTGGAGATGCTATAGTCAAGTTTCGCGACTACAACTCTGTAATGCAAGATCATCCAGTC 1800
Db 212961 TCTTGGAGATGCTATAGTCAAGTTTCGCGACTACAACTCTGTAATGCAAGATCATCCAGTC 212902
QY 1801 GAATCTCAAGCGAATAATAGAGGATCAGACAAAAGCTTACATCGGCAAGTGAACAAAGCC 1860
Db 212901 GAATCTCAAGCGAATAATAGAGGATCAGACAAAAGCTTACATCGGCAAGTGAACAAAGCC 212842
QY 1861 TCCACAGTTTGGCTATCCCTTATGTCAGTCTTAAATGACTCTACACAGAGTTTCATAGC 1920
Db 212841 TCCACAGTTTGGCTATCCCTTATGTCAGTCTTAAATGACTCTACACAGAGTTTCATAGC 212782
QY 1921 TAAATAGAAAGTTTGTGTTGTAAGGATCTAGGACAGCAGCTGAATATAAAGACATTTTC 1980
```

```

Db 212781 TAAATTAGAAAGTTGTTGCTGAAGATCTAGACAGCAGCTGAATAAAGCACATTC 212722
QY 1981 CTTTGAACGAACCTCTGTTTATTCAGCAGGCTCTGCTCAATCGCTCTCTATATTC 2040
Db 212721 CTTTGAACGAACCTCTGTTTATTCAGCAGGCTCTGCTCAATCGCTCTCTATATTC 212662
QY 2041 TGGTTATCTCCAAATACCAACACCTTAAGTGTTCGTTTGAGAGATTAATTATGCTGTTGGT 2100
Db 212661 TGGTTATCTCCAAATACCAACACCTTAAGTGTTCGTTTGAGAGATTAATTATGCTGTTGGT 212602
QY 2101 AAGCCCTTGTGAGGCTTACCAACACCTAGACACTAGACGATCTCAATAAATAAAGA 2156
Db 212601 AAGCCCTTGTGAGGCTTACCAACACCTAGACACTAGACGATCTCAATAAATAAAGA 212546

RESULT 5
AP002547/c
LOCUS 300550 bp DNA linear BCT 25-APR-2003
DEFINITION Chlamydothila pneumoniae J138 genomic DNA, complete sequence,
section 3/4.
ACCESSION AP002547 AB033786 AB033787 AB033816 AB033817 BA000008
VERSION AP002547.2 GI:10176693
KEYWORDS
SOURCE
ORGANISM Chlamydothila pneumoniae J138
Chlamydothila pneumoniae J138
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydothila.
REFERENCE
1 Shirai, M., Hirakawa, H., Ouchi, K., Tabuchi, M., Kishi, F., Kimoto, M.,
Takeuchi, A., Nishida, J., Shibata, K., Fujinaga, R., Yoneda, H.,
Matsushima, H., Tanaka, C., Furukawa, S., Miura, K., Nakazawa, A.,
Ishii, K., Shiota, T., Hattori, M., Kuhara, S. and Nakazawa, T.
Comparison of outer membrane protein genes omp and pmp in the whole
genome sequences of Chlamydia pneumoniae isolates from Japan and
the United States
J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)
2 J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)
JOURNAL MEDLINE
PUBMED 10839753
REFERENCE
2 Shirai, M., Hirakawa, H., Kimoto, M., Tabuchi, M., Kishi, F., Ouchi, K.,
Shiba, T., Ishii, K., Hattori, M., Kuhara, S. and Nakazawa, T.
Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA
Nucleic Acids Res. 28 (12), 2311-2314 (2000)
20330349
10871362
Shirai, M.
Direct Submission
Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
School of Medicine, Department of Microbiology; 1-1-1
Minamikogushi, Ube, Yamaguchi 755-8505, Japan
(E-mail: mshirai@po.cc.yamaguchi-u.ac.jp, Tel: 81-836-22-2227,
Fax: 81-836-22-2415)
On or before Sep 15, 2000 this sequence version replaced
gi:6172298, gi:6172300, gi:6172396, gi:6172398, gi:8978889.
FEATURES
Location/Qualifiers
1..300550
/organism="Chlamydothila pneumoniae J138"
/mol_type="genomic DNA"
/strain="J138"
/db_xref="taxon:138677"
/notes="synonym: Chlamydia pneumoniae (strain J138)"
complement (21..794)
/genes="dapF"
complement (21..794)
/genes="dapF"
/codon_start=1
/transl_table=11
/product="diaminopimelate epimerase"
/protein_id="BAA98725.1"
/db_xref="GI:8978890"
/translation="MAFVSPSISKXYFYSGAGNRLIGETLPEVEDVRLCQETRVDD
GFLYKPSCCADAQLIIFNSDGRPTMCGNGLRCATAHLASQKGSDDISVSTDSGLYS
GIFYSDRVLVDMTILADRASVHRLESREDPLPKEVCVCIHTGVPHAVWILPSTLIDL
SLIGPLRKHOTFSPDGVNPNFVLIHGLQCLRVTRYVERGVEGETAACGTCALASALVV
SNSYGMKESIQIHTMGELMTVSNQRGRVYLQGSVTRDL"
complement (763..1338)
/genes="clpP_1"
complement (763..1338)
/genes="clpP_1"
/codon_start=1
/transl_table=11
/product="CLP protease"
/protein_id="BAA98726.1"
/db_xref="GI:8978891"
/translation="WADGEVHLRDLIELLEEARVFFSEPVTEKASDAIKKLWYL
ELDKPKPIFVINSFGSDAGFAVMDQIKMLTSPVTVVTGLAASMGSVLSLCAAP
GRFATPHSRIMHQFSIGGPTGQATDLDIHAREILTKARIIDVYVEATNQPRDI
EKAIDRDMWTANEAKDFGLDGLILFSENDL"
complement (1358..2917)
/genes="glyA"
complement (1358..2917)
/genes="glyA"
/codon_start=1
/transl_table=11
/product="serine hydroxymethyltransferase"
/protein_id="BAA98727.1"
/db_xref="GI:8978892"
/translation="MLKVFEPKFAIVEIFTKVAVVSLHLKFLNAGSKKGQSLAS
TAYLAARHLNAPFSIGERIIDELKQSHRLMIASENSLSVLAMGNLLTKYK
ESGPAPRPFYSCCENVDIAIEWECVETAKELFAADCAQVPHSGADANLLAWAILTHKV
QCPAYKLGKTYKTNELTEETELKAEKSSCVGLGSLNSGHLTHGNVRLNVMKLM
RCFYDVNPDTCTFYAIRSLAKYKPKVLIAGYSSRRLNFAVLKQIAEDSCSVL
WYDMAHAGLVAGGVDEENPIPVADIVTTTHKTLPGRGGLVLAIREYESTLKA
CYLMMGGPLPHVIAKTAVALKEALSVDKFAHQVNNARLAEFLSHGLRLTGT
DNHNVVIDLGLSIGKTAEDLSVGIADVNNSLPSDAIGKWDTSGLRGLTPTTL
GMGIDMEEVADIIVKVLNRLSLCHVEGSSKKNKELPEAIAQEARDRVRLNLLRFP
LYPEIDLEALV"
3107..3778
/genes="CPj0522"
3107..3778
/genes="CPj0522"
/codon_start=1
/transl_table=11
/product="CT433 hypothetical protein"
/protein_id="BAA98728.1"
/db_xref="GI:8978893"
/translation="WTLLGLNQTKARKYQAHYLPILTFPYAKSTPONKRALQFLPQ
ATHVLTSPSTHLFLSRMTSLLSKATIKTKYLCIGESTKERLLSLFGQKVVVATQ
EAEGIFPLQALPSSARILYPHSLARPVIREFLYNRTFFSYPHYTVKPKLKNKI
LSKYKKIIFTSPITVRAFAKIFPRPEKTYWCQGRMTLQEFQKFSQKQVSLLETGK
SRTGP"
3843..4175
/genes="CPj0523"
3843..4175
/genes="CPj0523"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98729.1"
/db_xref="GI:8978894"
/translation="MASSATGFGDGTAPSLFPFPATPRYNFKLAFVTLIALVWIAL
IATTIAIGLCIHLPLCSFIITAIPLIYFISRYICSHYARNVYIALDVVDHSLQDMRS
HSPFSDR"
4223..5302
/genes="CPj0524"
4223..5302
/genes="CPj0524"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98730.1"
/db_xref="GI:8978895"
/translation="MSGFSPSTESSQVSLSVPRDKIAPKQFTIAKISTLIALSL
ALGAVAGISLITVILGNPFLALLITLALFSVWTVLVVHQTSKVSSNWKVLQNFQK

```

```

PLKAWQEKVNDVCSNEMQFNNHLPKFKVAICTDASOPFOPTFELTGLVIEKQST
GIIPNVGPNIIDNATNLSTILYSLKDSVWDCKREGGPAKGDFPSTFVYV
VKLPNALDQTNLNSAEEKKILFTFLGHVCPKSEELPNOQEYIRQALLAYENCL
KAAIBSHAVALPLFTSVYVEPPEILPKFEGTFYWDNQTAQCFKRALDAIQNTALR
YPORSLVLIQDPFNTIESQSRSEE"
5642..6406
/gene="CP10525"
CDS
5642..6406
/gene="CP10525"
/codon_start=1
/transl_table=11
/product="C7398 hypothetical protein"
/db_xref="GI:8978896"
/transl_table="MDALLSILAIQELDIKMLRVRVKEHOKELAKVQSLKSDIR
KVOBEKLEMLNLTQIRDSNGRIQEISEQIKLENOAAVKMDENALUTQEMTTANK
ERSLHQSDMLKQSGEDLIVSEKSLASTENSSVIEKEIFESIKKINEEGKAL
LEQRTBELKATNPILLISYIERLLNNKRDVVPIENRVCSGGCHVILTPQHENVLRKDD
RLIFCBHCSRILYQBSQVNAQENSTAKRRRRAAV"
complement(6827..7816)
/gene="kpsF"
CDS
complement(6827..7816)
/gene="kpsF"
/codon_start=1
/transl_table=11
/product="GutQ/kpsF family sugar-P isomerase"
/protein_id="BA98732.1"
/db_xref="GI:8978897"
/transl_table="MSPSMISTDVCQDILKQKEAVDFPQFPQKEAMQLAKILGH
SGWFFSGVGKSGCARKIVATLQISERALFFSPVDLLHGLDGLVSPGDIVCLFSKS
GETLELDTVPHLKSRALIVALTSMYPNSLAALSDLVILPVAELDPNLIPTNST
TCOMIFGDFLMLLPHRSGVSLSTYCKNHPSGOVGMKANGKVDNFKPKEVPPFCHLG
DKVFSLEVFSAVGCQVCIVDPQFRLMGIETFDGLRRSLASVGGVEVLSLEKMTA
NPKCITEDSDIALQLQMSSPVAVLPVLDNENRHVTGLLHMHTLAKAGLI"
complement(7843..9024)
/gene="sucB_2"
CDS
complement(7843..9024)
/gene="sucB_2"
/codon_start=1
/transl_table=11
/product="dihydrolipoamide succinyltransferase"
/protein_id="BA98733.1"
/db_xref="GI:8978898"
/transl_table="MIFEPFPKIGTSGGSIVRMLKMLGDHVARDEPLIEVSTDKI
ATELPSPKAGRLVFCVNEGVASDVLGLIELEBISEADDESTSCPILTSQETKSEA
GSSSSVWFSPVLSLAQREGIGLDNLQKTACTGKGRVTRQDLEAYISEQVSIPE
IPQGVNRPMPSPRLRAIASSLSKSDVPHASLVVDVDDVLDLNNLISGERQRFDDTH
GVKLITTFIVQCLAQTLKQFPPLNGSLDGTIVMKKSVNVGVAVNNLKEGVVVPVIH
NODRGLVSIKALADLSRARNKLDPSVEQDGSVTNFMGTGALIOMPIIRYPEV
AALGIGTIQKRVVVRDDSLAIRKMVYVTLTFDHRVLDLGIYGSEFLTSLNKRLSVTM
G"
complement(9038..10282)
/gene="glT"
CDS
complement(9038..10282)
/gene="glT"
/codon_start=1
/transl_table=11
/product="glutamate symport"
/protein_id="BA98734.1"
/db_xref="GI:8978899"
/transl_table="MKLMWKIFIGLVGVTLGLVLLEDKAIFPKPIGDIPLNLLSMVY
PLVFCSMVLGASISDMKKLIGIKSVGLYLGTTALAIIVIGLCFAMIFSPNGCDFA
QAQMDSAVTIDSNKTAAYFLSIIAQVFPSPNVRSPARGNIIQIIIFAILGIALRL
SEGRGRPVRFDIGFSEIMLRVMNMTMSPAPYGVGASMAWISGNHGLVWLQGLFII
AYLACLHATLVFGLVRFPGCKMSFSKLSMMDAISCAVSTASSATLPTRMCVS
KNLGSVAEVSQFVLPGATVNMNGTAIFQGMAAVFTAQYINCPILSLSLLLVWTATF
SAVGSAGVPGGMITLIGSVLASVGLPIQGIALLAGIDRLDIRLVGTPMNLGLAVVATY
VASGEGELSPYESIKQESVET"
complement(10282..11379)
/gene="ycaH"
CDS
complement(10282..11379)
/gene="ycaH"

```

```

Query Match 100.0%; Score 2156; DB 1; Length 300550;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTTAAAAACAGCTCGCATTAATATTATTAGTGAGAGCTTTTATTTTATTTT 60
Db 217698 ATAAATCTTTTAAAAACAGCTCGCATTAATATTATTAGTGAGAGCTTTTATTTTATTTT 217699

QY 61 ATAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 217638 ATAATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 217639

QY 121 AGTCTCTATAGACGAACAGACGACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 180
Db 217578 AGTCTCTATAGACGAACAGACGACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 217519

QY 181 GCGAGTGCAGCAATAAAGAGTGCAGAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC 240
Db 217518 GCGAGTGCAGCAATAAAGAGTGCAGAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC 217459

QY 241 TAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTTGCGTTCTGCAGTGAATGC 300
Db 217458 TAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTTGCGTTCTGCAGTGAATGC 217399

QY 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGTTCTAGTAACAGCTCGTCTTCTACTAG 360
Db 217398 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGTTCTAGTAACAGCTCGTCTTCTACTAG 217339

QY 361 CAGATCTGCAGACGTGGACTCAACGACGACGACCGCACCTACGCTCCCTCCACCCACGTT 420
Db 217338 CAGATCTGCAGACGTGGACTCAACGACGACGACCGCACCTACGCTCCCTCCACCCACGTT 217279

QY 421 TGATGATTAAGACTCAAGCGCAACAGCTTAACGATACTATCTTTACCTCAACATCACT 480
Db 217278 TGATGATTAAGACTCAAGCGCAACAGCTTAACGATACTATCTTTACCTCAACATCACT 217219

QY 481 AGCTGCACATACAGGCTGCTTTGCTGAGCTCCAGGATGCTGCTCACTAATAAAGATAC 540
Db 217218 AGCTGCACATACAGGCTGCTTTGCTGAGCTCCAGGATGCTGCTCACTAATAAAGATAC 217159

QY 541 AGCGGTACTGTAGAGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 600
Db 217158 AGCGGTACTGTAGAGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 217099

QY 601 AGTTAAAGTTGCGCGCAAAATACAGAAATACGAAATATGCTTGGATTAACCAAGCGAT 660
Db 217098 AGTTAAAGTTGCGCGCAAAATACAGAAATACGAAATATGCTTGGATTAACCAAGCGAT 217039

QY 661 TCTTGACTCTTTAGGTAAACTCACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 720
Db 217038 TCTTGACTCTTTAGGTAAACTCACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 216979

QY 721 TCTAGCAAACTAAACAAAGACGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTACT 780
Db 216978 TCTAGCAAACTAAACAAAGACGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTACT 216919

QY 781 CCCAGGAAAACGCTTGCATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 216918 CCCAGGAAAACGCTTGCATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 216859

QY 841 ACAGATAGAAAAGATGAAATGCGATTAGGATGCATATTTTTCAGGACAGAACGCTAG 900
Db 216858 ACAGATAGAAAAGATGAAATGCGATTAGGATGCATATTTTTCAGGACAGAACGCTAG 216799

QY 901 TCGAGCTGTAGAAAATGCTTAAATCTAATAACAGTATAAGCAACAATAGATTTCAGCTAAAGC 960
Db 216798 TCGAGCTGTAGAAAATGCTTAAATCTAATAACAGTATAAGCAACAATAGATTTCAGCTAAAGC 216739

QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGCATCTCC 1020
Db 216738 AGCAATCGCTACTGCTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGCATCTCC 216679

QY 1021 AATTCTTCAAGAACGGAACAAATCGTAAATACAGGCTGAGAAAGATCTTAAAAAATATCA 1080

```

Db 216678 AATCTTCAAGAGCGGAACAATGGTATACAGGCTGAGAAGATCTTAAAAATATCAA 216619
QY 1081 ACCTGCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAACAACA 1140
Db 216618 ACCTGCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAACAACA 216559
QY 1141 AGGAAGTAGTATGGTAGTATTCGGTTCCTGATGTTCCAGTCTGTTAGATGATGCTGAAATGAGAC 1200
Db 216558 AGGAAGTAGTATGGTAGTATTCGGTTCCTGATGTTCCAGTCTGTTAGATGATGCTGAAATGAGAC 216499
QY 1201 CGCTTCCATTTGATGCTCGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATATCC 1260
Db 216498 CGCTTCCATTTGATGCTCGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATATCC 216439
QY 1261 TGATTCCTAAGTTCGCCAACAGAGCTCCGACGACAAGCTAGAGCAGCGAAAGCGCTGG 1320
Db 216438 TGATTCCTAAGTTCGCCAACAGAGCTCCGACGACAAGCTAGAGCAGCGAAAGCGCTGG 216379
QY 1321 AGATGACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 1380
Db 216378 AGATGACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 216319
QY 1381 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 1440
Db 216318 TAAAGCTGGGCAACAAACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCT 216259
QY 1441 TGTGAGCGAGAGTTTCCTCCGCTCGCAGCAAGTTCATAGGGTCATCTGTAACAGCT 1500
Db 216258 TGTGAGCGAGAGTTTCCTCCGCTCGCAGCAAGTTCATAGGGTCATCTGTAACAGCT 216199
QY 1501 TTACAGAGCTCAAAATCTACAGGTTCTGATTAATAAACAAGATACAGAGTTATGA 1560
Db 216198 TTACAGAGCTCAAAATCTACAGGTTCTGATTAATAAACAAGATACAGAGTTATGA 216139
QY 1561 TGCTTCAAAATCATCAATGATGCTCTATGATGGGACGAAATGATCGACTCGTGAATG 1620
Db 216138 TGCTTCAAAATCATCAATGATGCTCTATGATGGGACGAAATGATCGACTCGTGAATG 216079
QY 1621 GATAACAATGTAAGTACCCCGCTCTCAGCATCCGTTCTTAGAGCAGCAGCAAGC 1680
Db 216078 GATAACAATGTAAGTACCCCGCTCTCAGCATCCGTTCTTAGAGCAGCAGCAAGC 216019
QY 1681 TCAGAGCAGCAAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC 1740
Db 216018 TCAGAGCAGCAAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAAC 215959
QY 1741 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTTACAAATCTGTAATCAGATCATCCAGTC 1800
Db 215958 TCTTGGAGATGCTATAGTCAAGTTTCGGCACTTACAAATCTGTAATCAGATCATCCAGTC 215899
QY 1801 GAATCTCAAGCGAATAAGTAGAGATCAGCAAAAGCTTACATCGGAGTGCAAGCC 1860
Db 215898 GAATCTCAAGCGAATAAGTAGAGATCAGCAAAAGCTTACATCGGAGTGCAAGCC 215839
QY 1861 TCACAGATTGGCTATCTTATGCAACTTTCTAATGACTCTTACAGAAAGTTCTATAGC 1920
Db 215838 TCACAGATTGGCTATCTTATGCAACTTTCTAATGACTCTTACAGAAAGTTCTATAGC 215779
QY 1921 TAAATTAGAAAGTTTGTGCTGAAGATCTAGGACAGAGCTGAATTAAGACATTTTC 1980
Db 215778 TAAATTAGAAAGTTTGTGCTGAAGATCTAGGACAGAGCTGAATTAAGACATTTTC 215719
QY 1981 CTTTGAACCAACTCTGTTTATTCAGCAGGTCCTGCTCAATATCGGCTCTCTATATTC 2040
Db 215718 CTTTGAACCAACTCTGTTTATTCAGCAGGTCCTGCTCAATATCGGCTCTCTATATTC 215659
QY 2041 TGGTTATCTCCAATAACAACACCTTAAGTTCGTTTGGAGAGATTAATATGCTTTGGT 2100
Db 215658 TGGTTATCTCCAATAACAACACCTTAAGTTCGTTTGGAGAGATTAATATGCTTTGGT 215599
QY 2101 AAGGCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 2156

Db 215598 AAGGCTTTGTTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 215543
RESULT 6
AR310754_08/c
WPCOMMENT
Sequence split into 13 fragments LOCUS AR310754 Accession AR310754
Fragment Name Begin End
AR310754_00 1 110000
AR310754_01 100001 210000
AR310754_02 200001 310000
AR310754_03 300001 410000
AR310754_04 400001 510000
AR310754_05 500001 610000
AR310754_06 600001 710000
AR310754_07 700001 810000
AR310754_08 800001 910000
AR310754_09 900001 1010000
AR310754_10 1000001 1110000
AR310754_11 1100001 1210000
AR310754_12 1200001 1230025
Continuation (9 of 13) of AR310754 from base 800001 (AR310754 Sequence 1 from patent US
Query Match 99.4%; Score 2143.4; DB 6; Length 110000;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2155; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 ATAAATCTTTAAACACAGGCTCGCATTAATTAATAGTGAGAGCTTTTATTTTTT 60
Db 28638 ATAAATCTTTAAACACAGGCTCGCATTAATTAATAGTGAGAGCTTTTATTTTTT 28639
QY 61 ATATAAACTAAAGATTTTATTTATTTTATGTTTTTATGTTTATGTTTATGTTTATGTTT 120
Db 28638 ATATAAACTAAAGATTTTATTTTATGTTTTTATGTTTATGTTTATGTTTATGTTT 28579
QY 121 AGGTCTATAGCAGAAACAGAACGCACACTCCCGCAGATCTTTCTGCTCAAGATTGA 180
Db 28578 AGGTCTATAGCAGAAACAGAACGCACACTCCCGCAGATCTTTCTGCTCAAGATTGA 28519
QY 181 GCGAGTGCAGCAATAAGAGTGGGAGCTCAAGATAGCAGGTGGGAGCTTAAGCC 240
Db 28518 GCGAGTGCAGCAATAAGAGTGGGAGCTCAAGATAGCAGGTGGGAGCTTAAGCC 28459
QY 241 TAAAGATCTTAAGACCGATTTCTGTAGAGCGATGGAGCATTTTGGCTTCTGCAAGTGAATGC 300
Db 28458 TAAAGATCTTAAGACCGATTTCTGTAGAGCGATGGAGCATTTTGGCTTCTGCAAGTGAATGC 28399
QY 301 TCTCATGATCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAACAGCTGCTTCTACTAG 360
Db 28398 TCTCATGATCTGGCAGATAAGCTGGGTATTTGCTTCTAGTAACAGCTGCTTCTACTAG 28339
QY 361 CAGATCTCAGACGTTGGACTCAACACAGCAGCCGACCTAGCCCTCTCCACCCACGTT 420
Db 28338 CAGATCTCAGACGTTGGACTCAACACAGCAGCCGACCTAGCCCTCTCCACCCACGTT 28279
QY 421 TCATGATTTATAGATCAAGCGCAACAGCTTACGATATATCTTCTTCTCAACATCACT 480
Db 28278 TCATGATTTATAGATCAAGCGCAACAGCTTACGATATATCTTCTTCTCAACATCACT 28219
QY 481 AGCTGACATACAGGCTGCTTTGTCAGCTCCAGGATCTGTCATTAATAAAGGATAC 540
Db 28218 AGCTGACATACAGGCTGCTTTGTCAGCTCCAGGATCTGTCATTAATAAAGGATAC 28159
QY 541 AGCGCTCTCTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACCTAAGAATCGCGATGC 600
Db 28158 AGCGCTCTCTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACCTAAGAATCGCGATGC 28099
QY 601 AGTTAAAGTTGGCGCAAAATPACAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 660
Db 28098 AGTTAAAGTTGGCGCAAAATPACAAATAGCGAAATATGCTTCGGATTAACCAAGCGAT 28039
QY 661 TCTTGACTCTTTAGGTAACTGACTTCTTCTGAGGCTTCTTCTTCTCAATC 720

Db 28038 TCCTGACTCTTTAGTAACTGACTCTCTCGACTCTTTACAGGCTGCTCTCTCCAAATC 27979
Qy 721 TGTAGCAAAACAAATACAAAGCAGCTGAGCTCTTTAAAGAGATGCAAGATACCCAGTAGT 780
Db 27978 TGTAGCAAAACAAATACAAAGCAGCTGAGCTCTTTAAAGAGATGCAAGATACCCAGTAGT 27919
Qy 781 CCCAGGAAAACGGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGAC 840
Db 27918 CCCAGGAAAACGGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGAC 27959
Qy 841 ACAGATAGAGAAAGATGGAATGCGATPAGGATGCAATATTTTCAGGACAGACGCTAG 900
Db 27858 ACAGATAGAGAAAGATGGAATGCGATPAGGATGCAATATTTTCAGGACAGACGCTAG 27799
Qy 901 TGGAGCTGTAGAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTCAGCTAAAGC 960
Db 27798 TGGAGCTGTAGAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTCAGCTAAAGC 27739
Qy 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCC 1020
Db 27738 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCC 27679
Qy 1021 AATTCTTCAAGAGCGGAACAAATGCTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1080
Db 27678 AATTCTTCAAGAGCGGAACAAATGCTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 27619
Qy 1081 ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACACAGTTGGAGGCTCCAAAGCAACA 1140
Db 27618 ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACACAGTTGGAGGCTCCAAAGCAACA 27559
Qy 1141 AGGAAGTAGTATTTGGTAGTATTTCTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
Db 27558 AGGAAGTAGTATTTGGTAGTATTTCTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 27499
Qy 1201 CGTTTCCATTTTGTATGCTGCGGTTTGGTTCAGATGATTCACATGTTCAATACGGAATATCC 1260
Db 27498 CGTTTCCATTTTGTATGCTGCGGTTTGGTTCAGATGATTCACATGTTCAATACGGAATATCC 27439
Qy 1261 TGATTTCAAGCTGCCAACAGGAGCTCGCAGCACCAAGCTAGCAGGAGGAAAGCGCTGG 1320
Db 27438 TGATTTCAAGCTGCCAACAGGAGCTCGCAGCACCAAGCTAGCAGGAGGAAAGCGCTGG 27379
Qy 1321 AGATGACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGG 1380
Db 27378 AGATGACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTTAGG 27319
Qy 1381 TAAAGCTGGGCAACAAAGGCGATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGTG- 1439
Db 27318 TAAAGCTGGGCAACAAAGGCGATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGTGT 27259
Qy 1440 TTGTGAGCGCAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGGTCATCTGTAAAAACAGC 1499
Db 27258 TTGTGAGCGCAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGGTCATCTGTAAACAGC 27199
Qy 1500 TTTTCAAGACCTCAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATG 1559
Db 27198 TTTTCAAGACCTCAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATG 27139
Qy 1560 ATGCTTACAAATCCATCAATGATGCTATGTTAGGGCAGAAATGATCGGACTCTGTATG 1619
Db 27138 ATGCTTACAAATCCATCAATGATGCTATGTTAGGGCAGAAATGATCGGACTCTGTATG 27079
Qy 1620 TGATTAACAATTAAGTACCCCGCTCTCACAGGATCCGTTCTTAGAGCAGCAACAGAGAG 1679
Db 27078 TGATTAACAATTAAGTACCCCGCTCTCACAGATCGTTCTTAGAGCAGCAACAGAGAG 27019
Qy 1680 CTCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGGTGTATTTCTGGCAATAGCAGAA 1739
Db 27018 CTCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGGTGTATTTCTGGCAATAGCAGAA 26959
Qy 1740 CTCCTGGAGATGCTTATAGTCAAGTTTTCGGCACTACATCTGTATGAGATCATCCAGT 1799
Db 26958 CTCCTGGAGATGCTTATAGTCAAGTTTTCGGCACTACATCTGTATGAGATCATCCAGT 26899

Qy 1800 CGAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTATCATCGGCAAGTCAAAAGC 1859
Db 26998 CGAATCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTATCATCGGCAAGTCAAAAGC 26839
Qy 1860 CTCACAGTTTGGCTATCCTTTATGTGCAACTTTCTTAATGACTCTACACAGAAGTTCAATAG 1919
Db 26838 CTCACAGTTTGGCTATCCTTTATGTGCAACTTTCTTAATGACTCTACACAGAAGTTCAATAG 26779
Qy 1920 CTAAATTTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAATAAAGCACTTT 1979
Db 26778 CTAAATTTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGCTGAATAAAGCACTTT 26719
Qy 1980 CCTTTGAAAACGAACTCCTCTGTTTATTCAGCAGGTCGTGCTCAATATCGGCTCTCTATATT 2039
Db 26718 CCTTTGAAAACGAACTCCTCTGTTTATTCAGCAGGTCGTGCTCAATATCGGCTCTCTATATT 26659
Qy 2040 CTGTTTATCTCCAAATAACAACACACTAAGTGTTCGTTTGGAGAGATTTATTTGCTTTGG 2099
Db 26658 CTGTTTATCTCCAAATAACAACACTAAGTGTTCGTTTGGAGAGATTTATTTGCTTTGG 26599
Qy 2100 TAAGGCCCTTTTGGAGGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 2156
Db 26598 TAAGGCCCTTTTGGAGGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 26542

RESULT 7
AX338293
LOCUS AX338293 1956 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 63 from Patent WO0181379.
ACCESSION AX338293
VERSION AX338293.1 GI:18128828
KEYWORDS
SOURCE Chlamydomophila pneumoniae
ORGANISM Chlamydomophila pneumoniae
REFERENCE 1
AUTHORS Bhatia, A., Probst, P. and Stromberg, E. J.
TITLE Compounds and methods for treatment and diagnosis of chlamydial infection
JOURNAL Patent: WO 0181379-A 63 01-NOV-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1. .1956
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"

ORIGIN
Query Match 90.7%; Score 1956; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGCTTAACTCTTATTTGCTCAGGTCCTATAGACAAAACAGACGACACCTCCGCGAGAT 160
Db 1 ATGCTTAACTCTTATTTGCTCAGGTCCTATAGACAAAACAGACGACACCTCCGCGAGAT 60
Qy 161 CTTTCTGCTCAAGGATTTGGAGGCGAGTGCAGCAATTAAGAGTGGGAAAGCTCAAGAATA 220
Db 61 CTTTCTGCTCAAGGATTTGGAGGCGAGTGCAGCAATTAAGAGTGGGAAAGCTCAAGAATA 120
Qy 221 GCAGGTGGGAAGCTAAGCCCTAAAGAACTTAAGACCGATTTCTGAGAGGATGAGGATC 280
Db 121 GCAGGTGGGAAGCTAAGCCCTAAAGAACTTAAGACCGATTTCTGAGAGGATGAGGATC 180
Qy 281 TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGCAGATAGCTGGGTATTTGCTTTAGT 340
Db 181 TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGCAGATAGCTGGGTATTTGCTTTAGT 240
Qy 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGCTGGACTCAACGACAGCGACCGACCT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGCTGGACTCAACGACAGCGACCGACCT 300

```
QY 401 ACGCTCTCCACCCACGTTTGTATATATAGACTCAAGCGCAAAACAGCTTACGATCT 460
Db 301 ACGCTCTCCACCCACGTTTGTATATATAGACTCAAGCGCAAAACAGCTTACGATCT 360
QY 461 ATCTTTACCTCAACATCACTAGCTGACATACAGCTGCTTTGGTGGAGCTCCAGATGCT 520
Db 361 ATCTTTACCTCAACATCACTAGCTGACATACAGCTGCTTTGGTGGAGCTCCAGATGCT 420
QY 521 GTCACTAATAAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 580
Db 421 GTCACTAATAAAGGATACAGCGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 480
QY 581 GAACTAAGATCCGATGCGAGTTAAAGTTGGCGGCAAAATACAGATTAAGCAATAT 640
Db 481 GAACTAAGATGCGGATGCGAGTTAAAGTTGGCGGCAAAATACAGATTAAGCAATAT 540
QY 641 GCTTCGGATAACCAAGGATCTTGGACTCTTTAGTTAACTGACTTCCTTCGACTCTTA 700
Db 541 GCTTCGGATAACCAAGGATCTTGGACTCTTTAGTTAACTGACTTCCTTCGACTCTTA 600
QY 701 CAGGCTGCTCTTCTCCAACTCTGTAGCAAAATACAAAGAGCTGAGCTTCTTAAAGAG 760
Db 601 CAGGCTGCTCTTCTCCAACTCTGTAGCAAAATACAAAGAGCTGAGCTTCTTAAAGAG 660
QY 761 ATGCAAGATAACCCAGTAGTCCCGAGGAAACGCTGCAATTTGCTCAATCTTTAGTTGAT 820
Db 661 ATGCAAGATAACCCAGTAGTCCCGAGGAAACGCTGCAATTTGCTCAATCTTTAGTTGAT 720
QY 821 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCTAT 880
Db 721 CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCTAT 780
QY 881 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTPATACAGATTAAGC 940
Db 781 TTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAATGCTAAATCTPATACAGATTAAGC 840
QY 941 AACATAGATTACGATTAAGCAGCAATCTACTGCTTAAGACACAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTACGATTAAGCAGCAATCTACTGCTTAAGACACAATAGCTGAAGCTCAG 900
QY 1001 AAAAAGTTCCCGACTCTCCAACTCTTCAAGAGCGGAACAAATGGTAAATCAGGCTCAG 1060
Db 901 AAAAAGTTCCCGACTCTCCAACTCTTCAAGAGCGGAACAAATGGTAAATCAGGCTCAG 960
QY 1061 AAAGATCTTTAAATAATCAAACTCGAGTGTCTGTAGTGTTCGAAATCCAGGAACTACA 1120
Db 961 AAAGATCTTTAAATAATCAAACTCGAGTGTCTGTAGTGTTCGAAATCCAGGAACTACA 1020
QY 1121 GTTGAGGCTCCAGCAACAAGGAAGTAGTATTGGTAGTATTCTGTGTTCATGCTGTTA 1180
Db 1021 GTTGAGGCTCCAGCAACAAGGAAGTAGTATTGGTAGTATTCTGTGTTCATGCTGTTA 1080
QY 1181 GATGATGCTGAAAATAGACCGCTTCCATTTTGTATGCTGTGGGTTTCGTCAGATGATTCAC 1240
Db 1081 GATGATGCTGAAAATAGACCGCTTCCATTTTGTATGCTGTGGGTTTCGTCAGATGATTCAC 1140
QY 1241 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCCACAGGAGCTCGCAGCAAGCT 1300
Db 1141 ATGTTCAATACGGAATCTGATTTCTCAAGCTGCCACAGGAGCTCGCAGCAAGCT 1200
QY 1301 AGAGCAGCGAAAGCGCTGGAGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAA 1360
Db 1201 AGAGCAGCGAAAGCGCTGGAGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAA 1260
QY 1361 GCTTTAGAACGGCTCTAGTTAAAGCTGGGCAACACAGGCGATCTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAACGGCTCTAGTTAAAGCTGGGCAACACAGGCGATCTCAATGCTTTAGGA 1320
QY 1421 CAGATCGCTTCTGCTGCTGTGTGAGCGCAGAGTTCTTCCTCCGCTGCGACAGATTTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGCTGTGTGAGCGCAGAGTTCTTCCTCCGCTGCGACAGATTTCTATA 1380
QY 1481 GGGTCATCTGTAAAAACAGCTTTTACAGACCTCAAAATCTACAGGTTCTGATTATAAAAA 1540
```

```
Db 1381 GGGTCATCTGTAAAAACAGCTTTTACAGACCTCAAAATCTACAGTTCTGATTATAAAAA 1440
QY 1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTTATGTTAGGCAACGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTTATGTTAGGCAACGA 1500
QY 1601 AATGATGCGACTGCTGATGATGATAAATGTAAGTACCCCGCTCTCACACGATCCGTT 1660
Db 1501 AATGATGCGACTGCTGATGATGATAAATGTAAGTACCCCGCTCTCACACGATCCGTT 1560
QY 1661 CCTAGACACAGAACAGAGCTCGAGGACCAAGAAAACAGATCAAGCCCTCGCTAGGGTG 1720
Db 1561 CCTAGACACAGAACAGAGCTCGAGGACCAAGAAAACAGATCAAGCCCTCGCTAGGGTG 1620
QY 1721 ATTTCTGGCAATAGCAAACTCTTGGAGATGTCTATAGTCAAGTTTTCGGCACTACAATCT 1780
Db 1621 ATTTCTGGCAATAGCAAACTCTTGGAGATGTCTATAGTCAAGTTTTCGGCACTACAATCT 1680
QY 1781 GTAATGAGATCATCCAGTCCGAATCCTCAAGCGAATTAATGAGGATCAGACAAAAGCTT 1840
Db 1681 GTAATGAGATCATCCAGTCCGAATCCTCAAGCGAATTAATGAGGATCAGACAAAAGCTT 1740
QY 1841 ACATCGGCACTGACAAAGCCTCCACAGTTTGGGTATCCTTATGTGCAACTTCTAATGAC 1900
Db 1741 ACATCGGCACTGACAAAGCCTCCACAGTTTGGGTATCCTTATGTGCAACTTCTAATGAC 1800
QY 1901 TCTACACAGAGTTCTAGCTAAATTAAGAAAGTTTGTGCTGAAGATCTAGGACAGCA 1960
Db 1801 TCTACACAGAGTTCTAGCTAAATTAAGAAAGTTTGTGCTGAAGATCTAGGACAGCA 1860
QY 1961 GCTCAAAATAAAGCACTTCTCTTGAACGAATCCTTGTGTTTATTTCAGCAGGTCCTGTC 2020
Db 1861 GCTCAAAATAAAGCACTTCTCTTGAACGAATCCTTGTGTTTATTTCAGCAGGTCCTGTC 1920
QY 2021 AATATCGGCTCTCTATATCTTGTGTTATCTCCAATAA 2056
Db 1921 AATATCGGCTCTCTATATCTTGTGTTATCTCCAATAA 1956
```

RESULT 8

AX349491
LOCUS AX349491 1956 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 14 from Patent WO202606.
ACCESSION AX349491
VERSION AX349491.1 GI:18615352
KEYWORDS

ORGANISM

Chlamydomonadales
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

REFERENCE

1
Ratti, G. and Grandi, G.
Immunisation against Chlamydia pneumoniae
Patent: WO 0202606-A 14 10-JAN-2002;
Chiron S.p.A. (IT)

FEATURES

source

1. 1956
/organism="Chlamydomonada pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"

ORIGIN

Query Match 90.7%; Score 1956; DB 6; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGTTAATCTTATGTTCCAGGTCCTATAGCAAGAAACAGACCACTCCCGAGAT 160
Db 1 ATGTTAATCTTATGTTCCAGGTCCTATAGCAAGAAACAGACCACTCCCGAGAT 60
QY 161 CTTTCTGCTCAAGATTTGAGGCGAGTGCAGCAATTAAGTGGGAGTCAAGATA 220
Db 61 CTTTCTGCTCAAGATTTGAGGCGAGTGCAGCAATTAAGTGGGAGTCAAGATA 120

221 GCAGGTGCGGAAGCTTAAGCTTAAGAAATCTAAGACCGATTCCTGTAGAGCGATGGAGCATC 280
Db GCAGGTGCGGAAGCTTAAGCTTAAGAAATCTAAGACCGATTCCTGTAGAGCGATGGAGCATC 180
281 TTGCGTTCTGCGAGTAATCTCTCATGAGTCTGGCAGATAAGCTGGTATTCCTCTAGT 340
Db TTGCGTTCTGCGAGTAATCTCTCATGAGTCTGGCAGATAAGCTGGTATTCCTCTAGT 240
341 AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGGACTCAACGACAGCGACCGACCT 400
Db AACAGCTCGTCTTCTACTAGCAGATCTGACAGCTGGACTCAACGACAGCGACCGACCT 300
401 ACSCCTCTCCACCCACGGTTTGAATTAATAGACTCAAGCGGCAACACAGCTTAAGATCT 460
Db ACSCCTCTCCACCCACGGTTTGAATTAATAGACTCAAGCGGCAACACAGCTTAAGATCT 360
461 ATCTTTACCTCAACATCAGCTAGCTGACATACAGCTGCTTTGGTGAAGCTTCCAGGATGCT 520
Db ATCTTTACCTCAACATCAGCTAGCTGACATACAGCTGCTTTGGTGAAGCTTCCAGGATGCT 420
521 GTCACATAATATAAGGATACAGCGGCTACTGTAGAGGAAACCGCAATCGCTCGGAGTGG 580
Db GTCACATAATATAAGGATACAGCGGCTACTGTAGAGGAAACCGCAATCGCTCGGAGTGG 480
581 GAAACTAAGATCCGATGCGATTAAGTTGGCGCGCAAAATACAGAAATTAGCGAATAT 640
Db GAAACTAAGATCCGATGCGATTAAGTTGGCGCGCAAAATACAGAAATTAGCGAATAT 540
641 GCTTCGGATAACCAAGGATCTTTGACTCTTTAGGTAACTGACTTCTTCGACCTCTTA 700
Db GCTTCGGATAACCAAGGATCTTTGACTCTTTAGGTAACTGACTTCTTCGACCTCTTA 600
701 CAGGCTGCTCTTCTTCCAACTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 760
Db CAGGCTGCTCTTCTTCCAACTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG 660
761 ATGCAAGATTAACCGAGTAGTCCAGGAAACCGCTGCAATTCCTCAATCTTTAGTTGAT 820
Db ATGCAAGATTAACCGAGTAGTCCAGGAAACCGCTGCAATTCCTCAATCTTTAGTTGAT 720
821 CAGACAGATCTTACAGCGACACAGATAGAGAAAGATGGAATGCGAATGAGGATGCAAT 880
Db CAGACAGATCTTACAGCGACACAGATAGAGAAAGATGGAATGCGAATGAGGATGCAAT 780
881 TTTGCGAGCAGAAACGCTAGTGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGC 940
Db TTTGCGAGCAGAAACGCTAGTGAGCTGTAGAAAATGCTAAATCTAATAACAGTAAAGC 840
941 AACATAGATTCAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 1000
Db AACATAGATTCAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAG 900
1001 AAAAAGTTCCCGGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAATACAGGCTGAG 1060
Db AAAAAGTTCCCGGACTCTCCAAATCTTCAAGAGCGGAACAAATGGTAATACAGGCTGAG 960
1061 AAGATCTTAAATAATCAACCTGCGAGATGGTTCTGATGTTCCAAATCCAGGAACATACA 1120
Db AAGATCTTAAATAATCAACCTGCGAGATGGTTCTGATGTTCCAAATCCAGGAACATACA 1020
1121 GTTGAGGCTCCAGACCAAGGAGTAGTATGCTAGTATTCGTGTTTCCATGCTGTTA 1180
Db GTTGAGGCTCCAGACCAAGGAGTAGTATGCTAGTATTCGTGTTTCCATGCTGTTA 1080
1181 GATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGCTCGGTTTCGTAGATGATTCAC 1240
Db GATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGCTCGGTTTCGTAGATGATTCAC 1140
1241 ATGTTCAATACGGAATTCCTGATTCCTCAAGCTGCCCAACAGGAGCTCGAGCAAGCT 1300
Db ATGTTCAATACGGAATTCCTGATTCCTCAAGCTGCCCAACAGGAGCTCGAGCAAGCT 1200

1301 AGACAGCGAAAGCCGCTGGAGATGACAGTGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1360
Db AGACAGCGAAAGCCGCTGGAGATGACAGTGTCTGCTGAGCGCTGGCAGATGCTCAGAAA 1260
1361 GCTTTAGAAAGCGGCTCTAGGTAAGCTGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1420
Db GCTTTAGAAAGCGGCTCTAGGTAAGCTGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1320
1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGTTCCTCCGCTGCAGCAGATCTATA 1480
Db CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGTTCCTCCGCTGCAGCAGATCTATA 1380
1481 GGGTCATCTGTAAAAACAGCTTTACAGAGCTCAAAATCTACAGGTTCTGATTAATAACA 1540
Db GGGTCATCTGTAAAAACAGCTTTACAGAGCTCAAAATCTACAGGTTCTGATTAATAACA 1440
1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGGTAGGCAAGA 1600
Db CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTATGGTAGGCAAGA 1500
1601 AATGATGCGACTCGTGATGTGATTAACATGTAAGTACCCGCTCTCACAGATCCGTT 1660
Db AATGATGCGACTCGTGATGTGATTAACATGTAAGTACCCGCTCTCACAGATCCGTT 1560
1661 CCTAGACACGAAACAGAGCTCGAGGACAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1720
Db CCTAGACACGAAACAGAGCTCGAGGACAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1620
1721 ATTCTGGCAATAGCAGAACTCTTGAGATGTCTATGATCAAGTTTGGGCACTACATCT 1780
Db ATTCTGGCAATAGCAGAACTCTTGAGATGTCTATGATCAAGTTTGGGCACTACATCT 1680
1781 GTAATGCAATCATCAGTCGAATCTCAAGCGAATATGAGGAGATCAGACAAAAGCTT 1840
Db GTAATGCAATCATCAGTCGAATCTCAAGCGAATATGAGGAGATCAGACAAAAGCTT 1740
1841 ACATCGGAGTCACAAAGCTCCACAGTTTGGCTATCTTATGTGCAATTTCTAATGAC 1900
Db ACATCGGAGTCACAAAGCTCCACAGTTTGGCTATCTTATGTGCAATTTCTAATGAC 1800
1901 TCTACACAGAGTTCTATAGCTAAATPAGAAAGTTTGTGCTGAGGATCTAGGACAGA 1960
Db TCTACACAGAGTTCTATAGCTAAATPAGAAAGTTTGTGCTGAGGATCTAGGACAGA 1860
1961 GCTGAAATAAAGCACTTCTTGAAGCAACTCTGTTTATTCAGCAGGTGCTGCTC 2020
Db GCTGAAATAAAGCACTTCTTGAAGCAACTCTGTTTATTCAGCAGGTGCTGCTC 1920
2021 AATATCGGCTCTCTATATCTGTTTATCTCCAATAA 2056
Db AATATCGGCTCTCTATATCTGTTTATCTCCAATAA 1956

RESULT 9
AX361992LOCUS 1956 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 385 from Patent WO0208267.

ACCESSION

AX361992

VERSION

AX361992.1

KEYWORDS

GI:18694447

SOURCE

Chlamydia pneumoniae

ORGANISM

Chlamydia pneumoniae

REFERENCE

1

AUTHORS

Fling, S.P., Skeiky, Y.A., Probst, P., and Bhatia, A.

TITLE

Compounds and methods for treatment and diagnosis of chlamydial infection

JOURNAL

Patent: WO 0208267-A 385 31-JAN-2002;

FEATURES

Location/Qualifiers

1..1956

/organism="Chlamydia pneumoniae"

/mol_type="unassigned DNA"

ORIGIN									
/db_xref="taxon:83558"									
Query Match 90.7%; Score 1956; DB 6; Length 1956;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	101	ATGGTAAATCTTATGGTCCAGGTCCTATAGACGAAACAGAAACGACACCTCCCGCAGAT	160						
DB	1	ATGGTAAATCTTATGGTCCAGGTCCTATAGACGAAACAGAAACGACACCTCCCGCAGAT	60						
QY	161	CTTTCTGCTCRAGGATGGAGCGGAGTCAGACABATAGAGTGGCGAAGCTCAAGAGATA	220						
DB	61	CTTTCTGCTCAAGGATGGAGCGGAGTCAGACABATAGAGTGGCGAAGCTCAAGAGATA	120						
QY	221	GCAGGTGCGGAAGCTAAGCCTAAAGAACTTAAGACCGGATTCGTAGAGCGATGGAGCATC	280						
DB	121	GCAGGTGCGGAAGCTAAGCCTAAAGAACTTAAGACCGGATTCGTAGAGCGATGGAGCATC	180						
QY	281	TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGT	340						
DB	181	TTGGGTTCTGAGTGAATGCTCTCATGAGTCTGGCAGATAAGCTGGGTATTTGCTTCTAGT	240						
QY	341	AACAGCTGCTCTTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACT	400						
DB	241	AACAGCTGCTCTTACTAGCAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACT	300						
QY	401	ACGCTCTCTCCACCGTTTGATGATTATAAGATCAAGCGCAACAGCTTACGATACT	460						
DB	301	ACGCTCTCTCCACCGTTTGATGATTATAAGATCAAGCGCAACAGCTTACGATACT	360						
QY	461	ATCTTTACTGACATCACTAGTGCATACAGGCTGCTTTGGTGAGCCTCCAGATGCT	520						
DB	361	ATCTTTACTGACATCACTAGTGCATACAGGCTGCTTTGGTGAGCCTCCAGATGCT	420						
QY	521	GTCACTAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGG	580						
DB	421	GTCACTAATATAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGG	480						
QY	581	GAAACTAAGATGCCGATGCGATTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATAT	640						
DB	481	GAAACTAAGATGCCGATGCGATTAAAGTTGGCGGCAAAATACAGAAATAGCGAAATAT	540						
QY	641	GCTTGGGATAACCAAGGATCTTGACTCTTTAGGTAAACTGACTTCTTCCAGCTCTTA	700						
DB	541	GCTTGGGATAACCAAGGATCTTGACTCTTTAGGTAAACTGACTTCTTCCAGCTCTTA	600						
QY	701	CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG	760						
DB	601	CAGGCTGCTCTTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAG	660						
QY	761	ATGCAAGATAACCGAGTAGTCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGAT	820						
DB	661	ATGCAAGATAACCGAGTAGTCCAGGGAACCGCTGCAATGCTCAATCTTTAGTTGAT	720						
QY	821	CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAAATGCGAATAGGGATGCATAT	880						
DB	721	CAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAAATGCGAATAGGGATGCATAT	780						
QY	881	TTTGCGGACAGAACCGCTAGTGGAGTGTAGAAAATGCTTAAATCTTAATAACAGTAAAGC	940						
DB	781	TTTGCGGACAGAACCGCTAGTGGAGTGTAGAAAATGCTTAAATCTTAATAACAGTAAAGC	840						
QY	941	AACATAGATTCAGCTAAGACGCAATCGCTACTGCTAAGACACAATAGCTGAGCTCAG	1000						
DB	841	AACATAGATTCAGCTAAGACGCAATCGCTACTGCTAAGACACAATAGCTGAGCTCAG	900						
QY	1001	AAAAAGTTCCTCCGACTCTCCAATTTCTCAAGAGCGGAACAAATGGTAAATACAGGCTGAG	1060						
DB	901	AAAAAGTTCCTCCGACTCTCCAATTTCTCAAGAGCGGAACAAATGGTAAATACAGGCTGAG	960						
QY	1061	AAAGATCTTAAATAATATCAAACTCGAGATGTTCTGATGTTCCCAATCCAGGAATACA	1120						
DB									

RESULT 10
AX045133
LOCUS
DEFINITION
ACCESSION
VERSION

AX045133
Sequence 3 from Patent WO0066739.
AX045133
AX045133.1 GI:11343732

linear

DNA

1852 bp

PAT 24-NOV-2000

KEYWORDS

SOURCE Chlamydomophila pneumoniae

ORGANISM Chlamydomophila pneumoniae

REFERENCE Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.

AUTHORS

TITLE Murdin, A.D., Oomen, R.P., Wang, J., and Dunn, P.

(Chlamydia) antigens and corresponding dna fragments and uses

JOURNAL

Patent: WO 0066739 A 3 09-NOV-2000;

Aventis Pasteur Limited (CA)

Location/Qualifiers

FEATURES

source

1..1852 /organism="Chlamydomophila pneumoniae"

/mol_type="unassigned DNA"

/db_xref="taxon:83558"

CDS

1..1752 /note="unnamed protein product"

/codon_start=1

/transl_table=11

/protein_id="CAC17378.1"

/db_xref="GI:11343733"

/db_xref="REMODEL: CAC17378"

/translation="MSLADKLGIASSNSSSTSRSDVSTTAFTPTPPPTFDYKT
 QAOAYDFTFTSTLADIALVLIQDAVTNIXDIADEETAFAAEWETKNADAVKY
 GAOITELAKYSDNOAILDLSGLKLTSPDLQALLOS VANNKAAELIKEMQDNPVP
 GKTPLAQSLVDOTDATABTOLEKQGNALRDYAFAGONASGAVENAKSNNSISNIDSAG
 AAIATAKTQIABAOKKFPDPILOEABQNVIAEKDLXNIPADGSDVPNPGTVGGS
 KOOGSSIGSIRVSMILLDDAEENETASILMSGERQMHMENTENPDSQAAQELAAQARA
 AKAGDDSAALADAAQALFAALGAKGQQQILNALQIASAAVSWAGVPPAAASSI
 GSVVKLYKTSKTSQDYKTOISAGYDAYKINDAYGRANDATRDVINNVSTPALTR
 SVPRAREPKEKTDQALRVISGNSRTLDGVYSOVSALOSVMOITOSNPQANNEI
 ROKLTSAVTKPQGPVPVQLSNDSTQKFIKLSLFAEGSRSTAIEIKALSFETNSLF
 IQQVLVNIHGSYSLQ"

ORIGIN

Query Match 85.8%; Score 1848.8; DB 6; Length 1852;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

305 ATGAGTCTGGCAGATAAGCTGGGTATTCCTTAGTAACAGCTCTCTTCTACTAGCAGA 364
 |||||
 1 ATGAGTCTGGCAGATAAGCTGGGTATTCCTTAGTAACAGCTCTCTTCTACTAGCAGA 60
 |||||
 365 TCTGACAGAGTGAATCAACGACGACGACGACCTAGCCCTCCACCACAGTTTGAT 424
 |||||
 61 TCTGACAGAGTGAATCAACGACGACGACGACCTAGCCCTCCACCACAGTTTGAT 120
 |||||
 425 GATTATAAGACTCAAGCGCAACAGGTTACGATATCTATCTTTACCTCAACATCACTAGCT 484
 |||||
 121 GATTATAAGACTCAAGCGCAACAGGTTACGATATCTATCTTTACCTCAACATCACTAGCT 180
 |||||
 485 GACATACAGGCTGCTTTGTGAGCTCCAGGATGCTGTCATCTAATATAAGGATACAGG 544
 |||||
 181 GACATACAGGCTGCTTTGTGAGCTCCAGGATGCTGTCATCTAATATAAGGATACAGG 240
 |||||
 545 GCTACTGATGAGAAACCCCAATCGCTGGGAGTGGAACTAAGAAATCCCGATGCGATT 604
 |||||
 241 GCTACTGATGAGAAACCCCAATCGCTGGGAGTGGAACTAAGAAATCCCGATGCGATT 300
 |||||
 605 AAAGTTGGCGCAAAATTAAGAATPAGCAAAATATGCTTCGGATAAACCAAGCGATTCTT 664
 |||||
 301 AAAGTTGGCGCAAAATTAAGAATPAGCAAAATATGCTTCGGATAAACCAAGCGATTCTT 360
 |||||
 665 GACTCTTTAGTAAAGTACTTCTTCGACCTCTTACAGGCTGCTCTCTTCAATCTGTA 724
 |||||
 361 GACTCTTTAGTAAAGTACTTCTTCGACCTCTTACAGGCTGCTCTCTTCAATCTGTA 420
 |||||
 725 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 784
 |||||
 421 GCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
 |||||
 785 GGGAAACGCTGCAATTCGCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAG 844
 |||||

481 GGGAAACGCTGCTCAATTCGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
 |||||
 845 ATAGAGAAAGATGGAATCGGATAGGATGCAATTTTTCAGGACAGAAAGCGTAGTGGGA 904
 |||||
 541 ATAGAGAAAGATGGAATCGGATAGGATGCAATTTTTCAGGACAGAAAGCGTAGTGGGA 600
 |||||
 905 GCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTTCAGCTTAAAGCAGCA 964
 |||||
 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTTCAGCTTAAAGCAGCA 660
 |||||
 965 ATGCTACTGCTTAAGACACAAAATAGCTGAAAGCTCAGAAAAGTTTCCCGGACTCTCCAAATT 1024
 |||||
 661 ATGCTACTGCTTAAGACACAAAATAGCTGAAAGCTCAGAAAAGTTTCCCGGACTCTCCAAATT 720
 |||||
 1025 CTTCAAGAAAGCGGAACAAATGGTAAATACAGGCTGAGAAAAGATCTTAAAAAATATCAAAACCT 1084
 |||||
 721 CTTCAAGAAAGCGGAACAAATGGTAAATACAGGCTGAGAAAAGATCTTAAAAAATATCAAAACCT 780
 |||||
 1085 GCAGATGGTTCCTGATGTTCCAAATCCAGGAACATACTAGTGGAGCTCCAAAGCAACAAGGA 1144
 |||||
 781 GCAGATGGTTCCTGATGTTCCAAATCCAGGAACATACTAGTGGAGCTCCAAAGCAACAAGGA 840
 |||||
 1145 AGTAGTATGGTAGTATTCGHTGTTTCCATGCTGTTAGTAGATGCTGAAAATAGACACGCT 1204
 |||||
 841 AGTAGTATGGTAGTATTCGHTGTTTCCATGCTGTTAGTAGATGCTGAAAATAGACACGCT 900
 |||||
 1205 TCCATTTTGTGTCGCTGGTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCCTGAT 1264
 |||||
 901 TCCATTTTGTGTCGCTGGTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCCTGAT 960
 |||||
 1265 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAAGCTAGACAGCAAGAACGCGCTGGAGAT 1324
 |||||
 961 TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAAGCTAGACAGCAAGAACGCGCTGGAGAT 1020
 |||||
 1325 GACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTCTAGGTAAA 1384
 |||||
 1021 GACAGTGTGCTGTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
 |||||
 1385 GCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1444
 |||||
 1081 GCTGGGCAACAAACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
 |||||
 1445 AGCCGACAGAGTTCCTCCGCTGACAGCAAGTTCTATAGGGTCAATCTGTAAAAACAGCTTTAC 1504
 |||||
 1141 AGCCGACAGAGTTCCTCCGCTGACAGCAAGTTCTATAGGGTCAATCTGTAAAAACAGCTTTAC 1200
 |||||
 1505 AAGACCTCAAAATCTACAGGTTCTGATTTAATAACACAGATATCAGCAGGTTATGATGCT 1564
 |||||
 1201 AAGACCTCAAAATCTACAGGTTCTGATTTAATAACACAGATATCAGCAGGTTATGATGCT 1260
 |||||
 1565 TACAAATCCATCAATGATGCTGCTATGGTAGGCAACGAAATGATGCGACTCGTGTGATGATA 1624
 |||||
 1261 TACAAATCCATCAATGATGCTGCTATGGTAGGCAACGAAATGATGCGACTCGTGTGATGATA 1320
 |||||
 1625 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGAGACGCAACAGAAAGCTCGA 1684
 |||||
 1321 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGAGACGCAACAGAAAGCTCGA 1380
 |||||
 1685 GGAACAGAAAAACAGATCAAGCCCTCGTAGGGTGAATTTCTGGCAATAGCAGACTCTT 1744
 |||||
 1381 GGAACAGAAAAACAGATCAAGCCCTCGTAGGGTGAATTTCTGGCAATAGCAGACTCTT 1440
 |||||
 1745 GGAGATGCTTAGTCAAGTTTCGGCACTCAATCTGTAATGCGAGATCATCCAGTCCGAAT 1804
 |||||
 1441 GGAGATGCTTAGTCAAGTTTCGGCACTCAATCTGTAATGCGAGATCACTCAGTCCGAAT 1500
 |||||
 1805 COTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCTTCA 1864
 |||||
 1501 COTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCTTCA 1560
 |||||
 1865 CAGTTTGGCTATCCTTTATGTCACATTTCTAATGACTCTACACAGAAAGTTTATAGCTAAA 1924
 |||||
 1561 CAGTTTGGCTATCCTTTATGTCACATTTCTAATGACTCTACACAGAAAGTTTATAGCTAAA 1620
 |||||

```

QY 1925 TTGAAAGTTGTTCTCTGAAGGATCTAGACAGCAGCTGAAATAAAGACATTTCTCTTT 1984
Db 1621 TTGAAAGTTGTTGTTGTAAGGATCTAGACAGCAGCTGAAATAAAGACATTTCTCTTT 1680
QY 1985 GAAACGAATCTCTGTTTATCTAGCAGGCTGCTGGTCAATATCGGCTCTCTATATCTGGT 2044
Db 1681 GAAACGAATCTCTGTTTATCTAGCAGGCTGCTGGTCAATATCGGCTCTCTATATCTGGT 1740
QY 2045 TATCTCAATAAACAACACCTAAGTGTCTGTTGGAGAGATTAATATGCTTTGGTAAGG 2104
Db 1741 TATCTCAATAAACAACACCTAAGTGTCTGTTGGAGAGATTAATATGCTTTGGTAAGG 1800
QY 2105 CCTTTGTTAGGCTTTACCAACACCTAGAACATCTTCAATAAATAAAGA 2156
Db 1801 CCTTTGTTAGGCTTTACCAACACCTAGAACATCTTCAATAAATAAAGA 1852

RESULT 11
LOCUS AX045137 2238 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 7 from Patent WO0066739.
ACCESSION AX045137
VERSION AX045137.1 GI:11343736
KEYWORDS
SOURCE
ORGANISM
Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
1
REFERENCE
Murdin, A.D., Omen, R.P., Wang, J. and Dunn, P.
AUTHORS i (Chlamydia) antigens and corresponding dna fragments and uses
TITLE thereof
JOURNAL Patent: WO 0066739-A 7 09-NOV-2000;
Aventis Pasteur Limited (CA)
FEATURES
source
1. 2238
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
766..2238
/notes="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC17330.1"
/db_xref="GI:11343737"
/db_xref="REMTREMBL.CAC17330"
/translation="MNPVIGPIDEPTPPADLSNQGLEASANKSABAQRIAGAE
AKPKESKTSVERMSILRSVNALMSLADKLGIASNSSSTSRSDVDSTTATPTP
PPFTFDYKQAGTAYDTIFTSLSLADIQALVSLQDAVTNIRKDTAATDETAIAEW
ETKNADAVKVGAGITELAKYASDNQALIDSLGKLSFDLQALLOSVAANNKAAEL
KEMQDNVPVPGKTPALAQSLVDOTDQATQIEKQDNRAIRDAYFAGQVAGAVENAKSN
NSISNIDSAKAAIATAKQIAEAKKPPDPSPILQEAQOMVIOAEKDLKNIKPADGSDV
PNPGTTVGSGQSGSSIGSRVSMLLDQENETASILMSGPROMIHMENTENPDSQAA
QVELAQARAKAAGDSDSAALADAKALEALGXKAGQQQGLINALGQIASAAVSA
GVLPVQVLMIRYQAYVEQKLISEEDLNSVDHHEHHH"

CDS
Query Match 68.0%; Score 1465.8; DB 6; Length 2238;
Best Local Similarity 99.8%; Pred. No. 1.1e-310;
Matches 1478; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATAAAACTTTAAACACAGGCTGCATTAATATTAGTGAGAGCTTTTTTATTTTTT 60
Db 666 ATAAAACTTTAAACACAGGCTGCATTAATATTAGTGAGAGCTTTTTTATTTTTT 725
QY 61 ATATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 726 ATATAAACTAAAGATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTT 785
QY 121 AGGTCCTATAGACGAAACAGAACCCACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 180
Db 786 AGGTCCTATAGACGAAACAGAACCCACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 845

```

```

QY 181 GCGAGTGCAGCAATAAAGAGTGCAGAAAGCTCAAAAGATAGCAGGTGCGGAAGCTAAGCC 240
Db 846 GCGAGTGCAGCAATAAAGAGTGCAGAAAGCTCAAAAGATAGCAGGTGCGGAAGCTAAGCC 905
QY 241 TAAAGAACTAAGACCGGATTTCTGTAGAGCGATGGAGCATCTTTGGCTTTGCGAGTGAATGC 300
Db 906 TAAAGAACTAAGACCGGATTTCTGTAGAGCGATGGAGCATCTTTGGCTTTGCGAGTGAATGC 965
QY 301 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTTCTAGTAAACAGCTGCTTTCTACTAG 360
Db 966 TCTCATGAGTCTGGCAGATAAGCTGGGTATTGCTTTCTAGTAAACAGCTGCTTTCTACTAG 1025
QY 361 CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCTCCTAGCCCTCTCCACCCACGTT 420
Db 1026 CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCTCCTAGCCCTCTCCACCCACGTT 1085
QY 421 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACCTATCTTTTCTCTCAATCACT 480
Db 1086 TGATGATTATAAGACTCAAGCGCAACAGCTTACGATACCTATCTTTTCTCTCAATCACT 1145
QY 481 AGCTGACATACAGGCTGCTTTGCTGAGCCTCCAGGATGCTGTCACTAATATAAAGATAC 540
Db 1146 AGCTGACATACAGGCTGCTTTGCTGAGCCTCCAGGATGCTGTCACTAATATAAAGATAC 1205
QY 541 AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 600
Db 1206 AGCGGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGC 1265
QY 601 AGTTAAAGTTGGCGGCAAAATACAGAAATAGCGGAAATATGCTTCGGATAACCAAGCGAT 660
Db 1266 AGTTAAAGTTGGCGGCAAAATACAGAAATAGCGGAAATATGCTTCGGATAACCAAGCGAT 1325
QY 661 TCTTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 720
Db 1326 TCTTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCAATC 1385
QY 721 TGTAGCAAACTAACAAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCAAGTAGT 780
Db 1386 TGTAGCAAACTAACAAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCAAGTAGT 1445
QY 781 CCCAGGAAACGCCCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db 1446 CCCAGGAAACGCCCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 1505
QY 841 ACAGATAGAGAAAGATGGAATGCGGATAGGATGCATATTTTGACGACAGAACCTAG 900
Db 1506 ACAGATAGAGAAAGATGGAATGCGGATAGGATGCATATTTTGACGACAGAACCTAG 1565
QY 901 TGGAGCTGTAGAAAATGCTTAATCTTAATACAGTATAAGCAACATAGATTAGCTAAAGC 960
Db 1566 TGGAGCTGTAGAAAATGCTTAATCTTAATACAGTATAAGCAACATAGATTAGCTAAAGC 1625
QY 961 AGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1020
Db 1626 AGCAATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCTCCGACTCTCC 1685
QY 1021 AATTCTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATATCAA 1080
Db 1686 AATTCTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAATAATATCAA 1745
QY 1081 ACCTGAGATGGTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACA 1140
Db 1746 ACCTGAGATGGTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACA 1805
QY 1141 AGAAGTGTATTTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAATCAGAC 1200
Db 1806 AGAAGTGTATTTGGTAGTATTGCTGTTTCCATGCTGTTAGATGATGCTGAAATCAGAC 1865
QY 1201 CGCTTCCATTTTGTATGCTCTGGGTTTTCGTGAGATGATTCATGTTCAATACGAAATATCC 1260
Db 1866 CGCTTCCATTTTGTATGCTCTGGGTTTTCGTGAGATGATTCATGTTCAATACGAAATATCC 1925
QY 1261 TGATTTCTCAAGTGGCCCAACAGGAGCTCGCAGCAACAGCTAGAGCGGAAAGCCCTGG 1320

```

Db 1926 TGATTCTCAAGCTGCCAACAGGAGCTCGCAGCAAGCTAGAGCAGCAAGCGCTGG 1985
Qy 1321 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
Db 1986 AGATGACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 2045
Qy 1381 TAAAGCTGGGCAACAAAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db 2046 TAAAGCTGGGCAACAAAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 2105
Qy 1441 TGTGAGCGCAGAGTTCCTCCGCTGCGAGCAAGTCTATAG 1481
Db 2106 TGTGAGCGCAGAG-TACTCCGCTGCGAGCAAGTCTATGG 2145

RESULT 12
CHT76KDA 2148 bp DNA linear BCT 17-MAR-1994
LOCUS Chlamydia pneumoniae 76 kDa protein gene, complete cds.
DEFINITION L23921
ACCESSION L23921.1 GI:435961
VERSION 76 kDa protein.
KEYWORDS Chlamydia pneumoniae
SOURCE Chlamydia pneumoniae
ORGANISM Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia phila.
REFERENCE 1 (bases 1 to 2148)
AUTHORS Perez Melgosa, M., Kuo, C.C. and Campbell, L.A.
TITLE Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-kilodalton protein containing a species-specific epitope
JOURNAL Infect. Immun. 62 (3), 880-886 (1994)
MEDLINE 94156481
PUBMED 7509320
COMMENT Original source text: Chlamydia pneumoniae DNA.
FEATURES
Location/Qualifiers
1..2148
/organism="Chlamydia pneumoniae"
/mol_type="genomic DNA"
/db_xref="taxon:83558"
1..2148
/codon_start=1
/transl_table=1
/product="76 kDa protein"
/protein_id="AAA23117.1"
/db_xref="GI:435962"

translation="MTKKHVAWVEGILNLPKOFFVKCSVVDWNTFVPSSTTEKA
ATNWKYKVCWQWLVGKHQVWINGQKPELYLYGAFLNPLAKATKTLINGENIA
WFIGTGLGRKAGDMSATVRYEYVEALSVEIDVSGIGRGNLLKFWAQIAANYDP
KEANGFTNYGFSALYMYGTDLSFRAYGSKPANDKLGSDFTFRKFLGHSAGF
KSNFKTFKRLALIISEFFFIYFNKTRFLFFELVNPFGPGPIDERTPTAD
LSAQGLPAAKSAEQRIGAEAKPKESKDSVERWSILRSVAVNLSLADKLGTA
SSNSSTSSADVSTATAPPPPTPDYKTOACTVDTFTSTSLADIQALVS
LQDAVNIKDTAATDETAIAEWETKNADAVKGAQITELAKVSNQALDSLGK
TDFDLQALLQSANNKAEALLKEMQDNVPVPPKPAIAQSLVDOTDQATQIEKO
ENAIRDAYAGQASNAEVAKSNNSISNDSKAAIATATKQIAEAKKFPDPSPILO
EAEQWVQAQKLNKIKPADGSDVPNPPTGGSKQOGSSIGTSIRVSMILLDDAENETA
SILMSGFPMHFNTPNSQAAQBLAAQAAKAGDDSAALADAKALEAL
GKAGQQQGIILNALQIINAAVSVAGVLPLOQVL"

ORIGIN

Query Match 67.6%; Score 1458; DB 1; Length 2148;
Best Local Similarity 99.9%; Pred. No. 5,9e-309;
Matches 1480; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy 1 ATAAATCTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTTT 60
Db 668 ATAAATCTTAAACAGGCTCGCATTAATATTAGTGAGAGCTTTTATTTTTT 727
Qy 61 ATAATAAACTAAAGATTTTATTTATTTTGTAGTTTAA-TGGTTAACTTATTCGTC 119
Db 728 ATAATAAACTAAAGATTTTATTTATTTTGTAGTTTATTTTGTATTCGTTAACTTATTCGTC 787

Qy 120 CAGTCTCTATAGACGAAACAGAACGACACTCCCGCAGATCTTTCTGTCTCAAGGATTGG 179
Db 788 CAGTCTCTATAGACGAAACAGAACGACACTCCCGCAGATCTTTCTGTCTCAAGGATTGG 847
Qy 180 AGCGAGTGCAGCAAAATAAGAGTGCAGAACTCAAAGAAATAGCAGGTGCGGAAGCTAAGC 239
Db 848 AGCGAGTGCAGCAAAATAAGAGTGCAGAACTCAAAGAAATAGCAGGTGCGGAAGCTAAGC 907
Qy 240 CTAAGGAATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGCCTTCTGAGTGAATG 299
Db 908 CTAAGGAATCTAAGACCGATTCTGTAGAGCGATGGAGCATCTTGCCTTCTGAGTGAATG 967
Qy 300 CTCTCATGAGTCTGGCAGATAAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTA 359
Db 968 CTCTCATGAGTCTGGCAGATAAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTA 1027
Qy 360 GCAGATCTGAGACGCTGGACTCAACGACGACCGCACTTACGCTCTCCACCCACGT 419
Db 1028 GCAGATCTGAGACGCTGGACTCAACGACGACCGCACTTACGCTCTCCACCCACGT 1087
Qy 420 TTGATGATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTTACCTCACATCAC 479
Db 1088 TTGATGATTATAGACTCAAGCGCAACAGCTTACGATCTATCTTTTACCTCACATCAC 1147
Qy 480 TAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGTCACTAATATAAGGATA 539
Db 1148 TAGCTGACATACAGGCTGCTTTGGTGAAGCTCCAGGATGCTGTCACTAATATAAGGATA 1207
Qy 540 CAGGGCTACTGATAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATG 599
Db 1208 CAGGGCTACTGATAGGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATGCCGATG 1267
Qy 600 CAGTTAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGCTTCGATTAACCAAGCGA 659
Db 1268 CAGTTAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGCTTCGATTAACCAAGCGA 1327
Qy 660 TTTCTGACTCTTTAGGTAAACTGACTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAA 719
Db 1328 TTTCTGACTCTTTAGGTAAACTGACTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAA 1387
Qy 720 CTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAG 779
Db 1388 CTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAG 1447
Qy 780 TCCAGGGAACCGCTGCAATTTGCTCAATTTAGTTGATGATCAGACAGATGCTACAGCGA 839
Db 1448 TCCAGGGAACCGCTGCAATTTGCTCAATTTAGTTGATGATCAGACAGATGCTACAGCGA 1507
Qy 840 CACAGATAGAGAAAGATGGAATGCGATTAGGATGATATTTTGGCAGGACAGAAACGCTA 899
Db 1508 CACAGATAGAGAAAGATGGAATGCGATTAGGATGATATTTTGGCAGGACAGAAACGCTA 1567
Qy 900 GTGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTAAAG 959
Db 1568 GTGAGCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTAAAG 1627
Qy 960 CAGCAATCGCTACTGCTTAAGACACAATAAGCTCAAGCTCAGAAAGAGTTTCCCGAGCTCTC 1019
Db 1628 CAGCAATCGCTACTGCTTAAGACACAATAAGCTCAGAAAGAGTTTCCCGAGCTCTC 1687
Qy 1020 CAATTTCTCAAGAGCGGAACAAATGGTAAATCAGGCTGAGAAAGATCTTTAAATAATCA 1079
Db 1688 CAATTTCTCAAGAGCGGAACAAATGGTAAATCAGGCTGAGAAAGATCTTTAAATAATCA 1747
Qy 1080 AACCTCAGATGCTTCTGATGTTTCCAAATCCAGGAACATAGTTGAGGCTCCAGCAAC 1139
Db 1748 AACCTCAGATGCTTCTGATGTTTCCAAATCCAGGAACATAGTTGAGGCTCCAGCAAC 1807
Qy 1140 AAGGAAGTATGATGCTAGTATTCTGTTTCCATGCTGTTAGATGATGCTGAAATGAGA 1199
Db 1808 AAGGAAGTATGATGCTAGTATTCTGTTTCCATGCTGTTAGTATGATGCTGAAATGAGA 1867
Qy 1200 CCGCTTCCATTTTGTATGCTGGGTTTCTGTCAGATGATTCACATGTTCAATPACGGAATC 1259

```
Db      1868  CCGCTTCATTTGATCTGGGTTTCGTGAGATGATTCATGTTCAATACGGAATC 1927
Qy      1260  CTGATTTCTAAGCTGCCCAACAGAGAGCTCGCAGCACCAAGCTAGAGCAGCGAAAGCCGCTG 1319
Db      1928  CTGATTTCTAAGCTGCCCAACAGAGAGCTCGCAGCACCAAGCTAGAGCAGCGAAAGCCGCTG 1987
Qy      1320  GAGATGACAGTGTCTGCGAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAG 1379
Db      1988  GAGATGACAGTGTCTGCGAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAG 2047
Qy      1380  GTAAGCTGGGCAACACAGGGCATCTCAATGCTTTAGACAGATCGCTTCTGCTCTG 1439
Db      2048  GTAAGCTGGGCAACACAGGGCATCTCAATGCTTTAGACAGATCGCTTCTGCTCTG 2107
Qy      1440  TTGTGAGCGGAGGATTCCTCCGCTGCAGCAAGTTCTATAG 1481
Db      2108  TTGTGAGCGGAGGAG-TCTCCGCTGCAGCAAGTTCTATAG 2148

RESULT 13
AX045135
LOCUS      1456 bp      DNA      linear      PAT 24-NOV-2000
DEFINITION      Sequence 5 from Patent WO0066739.
ACCESSION      AX045135
VERSION      AX045135.1 GI:11343734
KEYWORDS
SOURCE      Chlamydia pneumoniae
ORGANISM      Chlamydia pneumoniae
REFERENCE      1
AUTHORS      Murdin,A.D., Owen,R.P., Wang,J. and Dunn,P.
TITLE      i(chlamydia) antigens and corresponding dna fragments and uses
JOURNAL      Patent: WO 0066739-A 5 09-NOV-2000;
FEATURES      Location/Qualifiers
source      1..1456
             /organism="Chlamydia pneumoniae"
             /mol_type="unassigned DNA"
             /db_xref="taxon:83558"
             101..>1452
             /note="unnamed protein product"
             /codon_start=1
             /transl_table=11
             /protein_id="CAC17379.1"
             /db_xref="GI:11343735"
             /db_xref="REMBL:CAC17379"
             /translation="MVNPIGFPIDETPTPPADLSAQGLEASAAKSAEAOIRAGAE
             AKPKESKTSVERWSILRSVNALMSLADKLGIASSNSSSSTSRSDVDSTTATPTP
             PPTFDYKTOAGTAYDTIPTSLADIQALVSLQDAVTNIKDTAATDETAIAEAW
             ETKNDAVKVGAITELATKASDNQALLSLKLTSLDQALLQALLOSVAANNKAELL
             KEMQDNVPEKPTPAQSLVDQTDATQIEKQGNIRDAYFAGQASGAVENAKSN
             NSISNDSAKAALAKTQIAEAKKPPDSPILQEAQPMVIOAEKDKLNKIKPADGSDV
             PNFGTIVGGSKQSGSTSRVSMLLDAENETASILMSFGRIHMFNFPDSDQA
             QOELAAQARAAKAAGDSDAAALADAKALEAALGRAGQOQGLNALQIASAAVVA
             G"

ORIGIN
Query Match      67.5%; Score 1455; DB 6; Length 1456;
Best Local Similarity      100.0%; Pred. No. 2,8e-308;
Matches 1455; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1  ATAAATCTTTAAACAGCGCTCCGATTAATATTAGTGAGAGCTTTTATTTT 60
Db      1  ATAAATCTTTAAACAGCGCTCCGATTAATATTAGTGAGAGCTTTTATTTT 60
Qy      61  ATATAAACTAAAGATTTTATTTATTTTGTAGTTTATGTTTATCTATTGTC 120
Db      61  ATATAAACTAAAGATTTTATTTATTTTGTAGTTTATGTTTATCTATTGTC 120
Qy      121  AGTCTCTATAGACGAACAGACGACCTCCGAGATCTTCTGCTCAAGGATTGGA 180
Db      121  AGTCTCTATAGACGAACAGACGACCTCCGAGATCTTCTGCTCAAGGATTGGA 180
```

```

QY 1261 TGATTCTCAAGCTCCCAACAGGAGCTCGACGACAAAGCTAGACGACGAAAGCGCTGG 1320
Db 1261 TGATTCTCAAGCTCCCAACAGGAGCTCGACGACAAAGCTAGACGACGAAAGCGCTGG 1320
QY 1321 AGATGACAGTCTGCTCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
Db 1321 AGATGACAGTCTGCTCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGG 1380
QY 1381 TAAAGCTGGGCAACAAACAGGCGATACCTACTGCTTTAGGACAGATCGCTTCTGCTGCTG 1440
Db 1381 TAAAGCTGGGCAACAAACAGGCGATACCTACTGCTTTAGGACAGATCGCTTCTGCTGCTG 1440
QY 1441 TGTGAGCGCAGAGT 1455
Db 1441 TGTGAGCGCAGAGT 1455

RESULT 14
LOCUS AE016994 304769 bp DNA linear BCT 15-AUG-2003
DEFINITION Chlamydomophila caviae GPIC section 1 of 4 of the complete genome.
ACCESSION AE016994 AE015925
VERSION AE016994.1 GI:29834116
KEYWORDS
SOURCE Chlamydomophila caviae GPIC
ORGANISM Chlamydomophila caviae GPIC
REFERENCE 1 (bases 1 to 304769)
AUTHORS Read,T.D., Myers,G.S., Brunham,R.C., Nelson,W.C., Paulsen,I.T.,
Heidelberg,J., Holtzaple,E., Khouri,H., Fedorova,N.B., Carty,H.A.,
Unayam,L., Haft,D., Peterson,J., Beanan,M.J., White,O.,
Salzberg,S.L., Hsia,R.C., McClarty,G., Rank,R.G., Bavoil,P.M. and
Fraser,C.M.
TITLE Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
examining the role of niche-specific genes in the evolution of the
Chlamydiaceae
Nucleic Acids Res. 31 (8), 2134-2147 (2003)
PUBMED 22569155
12682364
REFERENCE 2 (bases 1 to 304769)
AUTHORS Read,T., Myers,G., Brunham,R., Nelson,W., Paulsen,I.,
Heidelberg,J., Holtzaple,E., Khouri,H., Fedorova,N., Carty,H.,
Unayam,L., Haft,D., Peterson,J., Beanan,M., White,O., Salzberg,S.,
Hsia,R.C., McClarty,G., Rank,R., Bavoil,P. and Fraser,C.
Direct Submission
Submitted (29-OCT-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1. 304769
/organism="Chlamydomophila caviae GPIC"
/mol_type="genomic DNA"
/isolate="GPIC"
/db_xref="taxon:227941"
57. .1052
/gene="hemB"
/locus_tag="CCA00001"
57. .1052
/gene="hemB"
/locus_tag="CCA00001"
/locus_tag="CCA00001"
/notes="similar to GB:L16784, SP:P33395, and PID:290384;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="delta-aminolevulinic acid dehydratase"
/protein_id="AAP04754.1"
/db_xref="GI:29834117"
/translation="MSSALILRRPRNRKTAARDIYAETALLPQDFICFFVKEGKN
ICEISILPGYRMSIDLLKRLCSLGAVALFPVPSDLKDAYGYSNPNKI
LCKSIYEVKAPDLPCVSDIALDPYTHGHGMDQGEVINDSEVIFGNATLHAE
MGADIVKPMMDGVRAHRAKLDQSGWTNTLILSYKVYASSLYNFRDALGSHLOS
GDKRNYQNPKNVLELCSLDEQEGADMIMKPAGLYLDVLRVRSNTLPLAAVQ
VSGEYALIAADKMGWLKEMAFYESLIAIKRAGADMIIISVATLILEMLVSSKF"

```

gene

CDS

```

complement(1068..2480)
/gene="ndrA"
/locus_tag="CCA00002"
complement(1068..2480)
/gene="ndrA"
/locus_tag="CCA00002"
/notes="similar to SP:Q56586, GB:L16784, SP:P33395, and
PID:290384; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, subunit A"
/protein_id="AAP04755.1"
/db_xref="GI:29834118"
/translation="MKAIATRGDLDSLQSPKESGFKRIDPALVSVJLRPYSALALK
LKYQDDVSSGSPIAEYKFGVFITSPVSGIVKEIRRGESRLDVIWKTPGQNL
TEYSDLSLRAELLEIKFKGLFALFKQRPDIIPALPTQPDVDFINLADNRPTTP
STEKOLAVFSSREGEVFNVGVRVAKLFGCLPHIVSTRLALPKDLKSLIAHLKI
TGYPBSGSPSTHIIYIARTINEKDIIVTILSFOELVJIGHLPLKGRILNEQVVALAGSG
LKSLRRYVITTKANFESILLPQDISSDVLSLGDPLTGRICDRESLPCGWRDSTI
SVLPNPKTQAFNFKLGINPKPTHTTYLSGFLKRRKTYMDPDNTHGETRPLIDTEI
YDKMPLKIPVPLIKAVITKDFELACMLGFLVCEPDEFALPTFIDPSKTEMLTIKD
ALKDYAKETGILNLEYEDKE"
complement(2502..2936)
/locus_tag="CCA00003"
complement(2502..2936)
/locus_tag="CCA00003"
/locus_tag="CCA00003"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAP04756.1"
/db_xref="GI:29834119"
/translation="MNSNSAQKILDSIKOILTLYNIDFDFSGSSSSDSEVYEVYIM
ETQKEIKELDKESQDILOQTGTREQMEVFANNPNFSPPEWLALENIRSSDAYKK
ETPALNIEVTKLGGTETKSSNNKTKTSSKKKKKWIPL"
3048..5201
/gene="greA"
/locus_tag="CCA00004"
3048..5201
/gene="greA"
/locus_tag="CCA00004"
/notes="similar to GB:L10328, SP:P04846, GB:M12163,
PID:146649, PID:290510, GB:U00096, and PID:1790093;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="transcription elongation factor GreA"
/protein_id="AAP04757.1"
/db_xref="GI:29834120"
/translation="MDYLEKLQVLIDEEQSPFFNWEYCFNDVVRGAELVQILEKV
KHSLAPLFGKIADTVLPWEMIPDGKEDRVQLVDVONTNAKPFYDAADIVNRK
YAGRNFEALRVGLRDGREFQYLSRFDLFMHLSEGNFVHSGGVMGVSVPFL
QQKVLIEFGVMMKADISPEAFKSLIPLEKDFLSRRFDGPDGPFESYAKEHPSVIE
MLLKDLGPKTAKEDLDVLPVPEAMNRWQATKTKKNCITTPKTIKDYRYN
PGGSLITOLETFLSOIETDSAKLVEIYQFIRDLHSELKKNPREIVIVLQKAFMLIKK
NQSLETORLLISDFLGEKSVLDSEFLSPEEIVSVNNISIVVLQKAFMLIKK
HSPWENVFMKFLSTTSPSLRELTFKVFGEDSYRKKLEKLECAEHPMTPPEVFT
WFFELGSHDDGIFKPNDRKIERLFEALVFMVHASTPQKELGKIVTFLVSQRYL
AVRNIEGAPLSYKEILLISTKCSQSFSSDLSVLQSLAEVHPDLKKNITVVVEEDIL
WTPPESTKMNKLQSLSGKGMVNAKETEDARVLGDLRENSYKFALEKRLQBEI
RVLSEVNRARILTKDIVTDKVGCGKYSLEDNGLNLTYSILGFWDDADPDNVLISL
KSKLAQEMLDKSVGAAVQGGKKYKISRISQSIWDE"
5353..6543
/locus_tag="CCA00005"
5353..6543
/locus_tag="CCA00005"
/notes="similar to SP:P04693, GB:L10328, SP:P04846,
GB:M12163, PID:146649, PID:290510, GB:U00096, and
PID:1790093; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="amino acid biosynthesis aminotransferase, class
I"

```

gene

CDS

RESULT 15	BD264444	1537 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD264444				
DEFINITION	Compounds and methods for treatment and diagnosis of chlamydial infection.				
ACCESSION	BD264444				
VERSION	BD264444.1	GI:33074212			
KEYWORDS	JP 2002531129-A/32..				
SOURCE	Chlamydia trachomatis				
ORGANISM	Chlamydia trachomatis				
REFERENCE	1 (bases 1 to 1537)				
AUTHORS	Probst,P., Bhatia,A., Skeiky,Y.A.W., Fling,S.P., Jen,S. and Stromberg,E.J.				
TITLE	Compounds and methods for treatment and diagnosis of chlamydial infection				
JOURNAL	Patent: JP 2002531129-A 32 24-SEP-2002;				
COMMENT	CORIXA CORP				
	OS Chlamydia trachomatis				
	PN JP 2002531129-A/32				
	PD 24-SEP-2002				
	PF 08-DEC-1999 JP 2000586916				
	PR 08-DEC-1998 US 09/208277,08-APR-1999 US 09/288594 PR				
	01-OCT-1999 US 09/410568,22-OCT-1999 US 09/426571 PI PETER				
	PROBST, AJAY BHATIA, YASIR A W SKEIKY, STEVEN P FLING, PI SHYIAN				
	JEN,				
	PI ERICA JEAN STROMBERG				
	PC C12N15/09,C12N15/09,A61K35/12,A61K38/00,A61K39/118,A61K39/395,				
	PC A61K39/395,				
	PC A61K45/00,A61K48/00,A61P33/00,C07K14/295,C07K16/12,C07K19/00,				
	PC C12N1/15,				
	PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/532,				
	PC G01N33/569//(C12N1/19,C12R1/01),(C12N15/09,C12R1/01),(C12Q1/68, PC				
	C12R1/01),				
	PC C12N15/00,C12N15/00,C12N5/00,A61K37/02,(C12N15/00,C12R1/01) CC				
	Compounds and methods for treatment and diagnosis of CC				
	chlamydial infection				
FH Key	Location/Qualifiers				
FT source	1..1537				
FT	/organism='Chlamydia trachomatis'.				
FEATURES	Location/Qualifiers				
source	1..1537				
	/organism="Chlamydia trachomatis"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:813"				
ORIGIN					
Query Match	8.7%;	Score 187.2;	DB 6;	Length 1537;	
Best Local Similarity	53.0%;	Prsd. No. 2.7e-30;			
Matches 480;	Conservative	0;	Mismatches 408;	Indels 18;	Gaps 3;
QY	1167	TTTCCATGCTGTAGATGATGCTGAAATAGACCGCTTCATTTGATGCTCGGGTTTC	1226		
Db	186	TTTCCATGCTGTAGATGATGCTGAAATAGACCGCTTCATTTGATGCTCGGGTTTC	245		
QY	1227	GTCCAGATGATTCACATGTTCAAACGGAAATCCTGATTCCTCAAGCTGCCAACAGGAGC	1286		
Db	246	GATCTATGATCGAACAAATTAATGTAACAATCTCTGCAACAGCTAAAGAGCTACAGCTA	305		
QY	1287	TGCGACGACAAAGCTAGACGACGAAAGCCGCTGGAGATGACATGCTGTGCGAGCGCTGG	1346		
Db	306	TGGAGGCTCAGCTGA---CTGCGATGTCAGATCAACTGTTGGTGGGATGGCGAGCTCC	362		
QY	1347	CAGATGCTCAGAAAGCTTTAGAGCGCTCTAGTAAAGCTGGGCAACAC	1397		
Db	363	CAGCCGAATACAGCAATCAAGATGCTCTTCCGCAAGCTTTTGAACAACACATCAGCAG	422		
QY	1398	AGGGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGTTC	1457		

Db 423 ATGGTTTAGCTACAGCTATGGCAAGTGGCTTTTGAGCTGCCAAGGTTGGAGAGGCT 482
QY 1458 CTCCTGGCTGAGCAAGTTCTATAGGTCATCTGTAAAAAGAGCTTTTACAGACCTCAAAAT 1517
Db 483 CCGCAGCAACAGCTGGCACTGTCCAGATGAATGTAACACAGCTTTTACAGACAGCGTTT 542
QY 1518 CTACAGGTTCT-----GATTATAAAACACAGATATCAGCAGGTTATGATGCTTACAAAT 1571
Db 543 CTTGCACTTCTTCAGCTCTTTATGCGACGACCTTCCGATGGATATTCGCTTACAAA 602
QY 1572 CCATCAATGATGCTATGGTAGGCGACGAATGATCGACTCGTGATGTGATAAAACAATG 1631
Db 603 CACTGAACCTTTATATTCCGAAAGCAGAGCGCGTGCAGTCAGCTATTAGTCAAAC 662
QY 1632 TAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGCTCGAGGACCCAG 1691
Db 663 CAAATCCCGCTTTCCAGAAAGCGTTTCTCGTTCTGGCATAGAAAGTCAAGGACGCGAGTG 722
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTTTGGAGATG 1751
Db 723 CAGATGCTAGCCAAAGAGCAGAGAACTATTGTCAGAGATAGCCAAACGTTAGGTGATG 782
QY 1752 TCTATAGTCAAGTTTCGGCACTCAATCTGTAAATGAGATCATCCAGTCAATCCTCAAG 1811
Db 783 TATATAGCCGCTTACAGGTTCTGGATTCTTTGATGTCTAGATTGTGAGCAATCCGCAAG 842
QY 1812 CGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCCCTCCACAGTTTG 1871
Db 843 CAAATCAAGAGAGATTATCGAGAGCTACGGCATCTATTAGCAAGCTCCCAATTTG 902
QY 1872 GCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAATTAGAAA 1931
Db 903 GGTATCCTGCTGTTTCAAGATTCTGTGATAGCTTGCAGAGTTTGTGCACANTTGGAAA 962
QY 1932 GTTTGTTTGTGAGGATCTAGGACAGAGCTGAAATATAAGCACTTTCCTTTGAAACGA 1991
Db 963 GAGAGTTTGTGATGGGAAACGTAGTCTCGCAGAAATCTCAAGAGAAATGCGTTTAGAAAAC 1022
QY 1992 ACCTCTTTGTTTATTTCAGCAGGCTGCTCAATATCGGCTCTCTATATCTGTTATCTCC 2051
Db 1023 AGCCCGTTTCATTCACAGGTTGTTGTAACATTCGCTTCTTATCTCTGTTATCTTT 1082
QY 2052 AATAAC 2057
Db 1083 CTTAAC 1088

Search completed: March 24, 2004, 01:57:06
Job time : 8390.96 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	2156	100.0	2156	10	US-09-564-479-1	Sequence 1, Appli
c	2143.4	99.4	1230025	15	US-10-289-763-1	Sequence 1, Appli
2	1956	90.7	1956	9	US-09-841-132-385	Sequence 385, App
3	1956	90.7	1956	9	US-09-841-132-385	Sequence 63, Appl
4	1956	90.7	1956	13	US-10-007-693-63	Sequence 63, Appl
5	1956	90.7	1956	13	US-10-007-693-63	Sequence 14, Appl
6	1956	90.7	1956	13	US-10-012-273-14	Sequence 14, Appl
7	1848.8	85.8	1852	10	US-09-564-479-3	Sequence 3, Appli
8	1465.8	68.0	2238	10	US-09-564-479-7	Sequence 7, Appli
9	1455	67.5	1456	10	US-09-564-479-5	Sequence 5, Appli
10	187.2	8.7	1537	9	US-09-841-132-57	Sequence 57, Appl
11	185.4	8.6	1941	9	US-09-841-132-412	Sequence 412, App
12	184	8.5	1171	10	US-09-841-260-13	Sequence 13, Appl
13	184	8.5	1171	13	US-10-007-693-13	Sequence 13, Appl
14	184	8.5	1834	10	US-09-841-260-38	Sequence 38, Appl
15	184	8.5	1834	13	US-10-007-693-38	Sequence 38, Appl

Qy	Qy	Db	Qy
1	1	1	61
A T A A A A C T T T T A A A A A C A G G C T C G C A T T A A T T A T T A G T G A G A G C T T T T T T T A T T T T T	A T A A A A C T T T T A A A A A C A G G C T C G C A T T A A T T A T T A G T G A G A G C T T T T T T T A T T T T T	A T A A A A C T T T T A A A A A C A G G C T C G C A T T A A T T A T T A G T G A G A G C T T T T T T T A T T T T T	A T A A T A A A C T A A A A A G A T T T T T A T T T T T T T T T T T T T T T T A T T A G T T A A T C C T A T T T G G T C C
60	60	60	120

;; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
;; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
;; TITLE OF INVENTION: and treatment of infection
;; FILE REFERENCE: 9710-003-999
;; CURRENT APPLICATION NUMBER: US/10/289,762
;; CURRENT FILING DATE: 2003-03-27
;; NUMBER OF SEQ ID NOS: 6849
;; SEQ ID NO 1
;; LENGTH: 1230025
;; TYPE: DNA
;; ORGANISM: Chlamydia pneumoniae
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)-(15000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (15001)-(30000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (30001)-(45000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (45001)-(60000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (60001)-(75000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (75001)-(90000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (90001)-(105000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (105001)-(120000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (120001)-(135000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (135001)-(150000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (150001)-(165000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (165001)-(180000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (180001)-(195000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (195001)-(210000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (210001)-(225000)
;; OTHER INFORMATION: n=a or c or g or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (225001)-(240000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (240001)-(255000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (255001)-(270000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (270001)-(285000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (285001)-(300000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (300001)-(315000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (315001)-(330000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (330001)-(345000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (345001)-(360000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (360001)-(375000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (375001)-(390000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (390001)-(405000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (405001)-(420000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (420001)-(435000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (435001)-(450000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (450001)-(465000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (465001)-(480000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (480001)-(495000)
;; OTHER INFORMATION: n=a or c or g or t
;; NAME/KEY: misc feature
;; LOCATION: (495001)-(510000)

/	OTHER INFORMATION: n=a or c or g or t	Db	828458	TAAAGATCTAAGACCGATTCTGTAGACGATGGAGCATCTTGGCTTCGCAATGATGC	828399
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	301	TCTCATGTGCTGGCAGATAAGCTGGGTATTCCTCTTAGTAAACAGCTGCTCTTACTAG	360
/	LOCATION: (510001)..(525000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828398	TCTCATGTGCTGGCAGATAAGCTGGGTATTCCTCTTAGTAAACAGCTGCTCTTACTAG	828339
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	361	CAGATCTGCAGACGTGGACTCAACGACGACCGACCTAGCGCTCTCCACCCACGTT	420
/	LOCATION: (525001)..(540000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828338	CAGATCTGCAGACGTGGACTCAACGACGACCGACCTAGCGCTCTCCACCCACGTT	828279
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	421	TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTACTTCTTACCTCAACATCACT	480
/	LOCATION: (540001)..(555000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828278	TGATGATTATAGACTCAAGCGCAACAGCTTACGATCTACTTCTTACCTCAACATCACT	828219
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	481	AGCTGACATACAGGCTGCTTGGTGGAGCTCCAGATGCTGTCACTAATATAAAGATAC	540
/	LOCATION: (555001)..(570000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828218	AGCTGACATACAGGCTGCTTGGTGGAGCTCCAGATGCTGTCACTAATATAAAGATAC	828159
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	541	AGCGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAATGCCGATGC	600
/	LOCATION: (570001)..(585000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828158	AGCGCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAATGCCGATGC	828099
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	601	AGTTAAAGTTGGCGCGCAAAATTACAGAATTACGAAATATGCTTCGGATAACCAAGCGAT	660
/	LOCATION: (585001)..(600000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828098	AGTTAAAGTTGGCGCGCAAAATTACAGAATTACGAAATATGCTTCGGATAACCAAGCGAT	828039
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	661	TCTTGACTCTTTAGGTAACTGACTTCTTGGAGCTCTTACAGGCTGCTCTTCTCAATC	720
/	LOCATION: (600001)..(615000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	828038	TCTTGACTCTTTAGGTAACTGACTTCTTGGAGCTCTTACAGGCTGCTCTTCTCAATC	827979
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	721	TGTAGCAACCAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	780
/	LOCATION: (615001)..(630000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827978	TGTAGCAACCAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT	827919
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	781	CCCAGGAAACCGCTGCAATTGCTCAATCTTTAGTTGATTCAGACAGATGCTTACAGCGAC	840
/	LOCATION: (630001)..(645000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827918	CCCAGGAAACCGCTGCAATTGCTCAATCTTTAGTTGATTCAGACAGATGCTTACAGCGAC	827859
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	841	ACAGATAGAGAAAGATGCAAAATGCGATTAGGAGTGCATATTTTGCAGACAGACGCTAG	900
/	LOCATION: (645001)..(660000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827858	ACAGATAGAGAAAGATGCAAAATGCGATTAGGAGTGCATATTTTGCAGACAGACGCTAG	827799
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	901	TGGAGCTGTAGAAAATGTCTAAATCTAATAACAGTATAAGCAACATAGATTCAAGCTAAC	960
/	LOCATION: (660001)..(675000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827798	TGGAGCTGTAGAAAATGTCTAAATCTAATAACAGTATAAGCAACATAGATTCAAGCTAAC	827739
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	961	AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC	1020
/	LOCATION: (675001)..(690000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827738	AGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCC	827679
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1021	AAATCTTCAAGAGCGGAAACAAATGCGTAATACAGGCTCAGAAAAAGATCTTAAATATCAA	1080
/	LOCATION: (690001)..(705000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827678	AAATCTTCAAGAGCGGAAACAAATGCGTAATACAGGCTCAGAAAAAGATCTTAAATATCAA	827619
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1081	ACCTCAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGAACACA	1140
/	LOCATION: (705001)..(720000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827618	ACCTCAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGAACACA	827559
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1141	AGGAAGTAGTATGTTAGTATTCGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC	1200
/	LOCATION: (720001)..(735000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827558	AGGAAGTAGTATGTTAGTATTCGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC	827499
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1201	CGCTTCCATTTTGTGCTGCTGTTTTCGTCAGATGATTCACATGTTCAATACGGAATATCC	1260
/	LOCATION: (735001)..(750000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827498	CGCTTCCATTTTGTGCTGCTGTTTTCGTCAGATGATTCACATGTTCAATACGGAATATCC	827439
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1261	TGATTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
/	LOCATION: (750001)..(765000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827438	TGATTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	827379
/	FEATURE:				
/	NAME/KEY: misc_feature	QY	1321	AGATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1380
/	LOCATION: (765001)..(780000)				
/	OTHER INFORMATION: n=a or c or g or t	Db	827378	AGATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	827319
/	FEATURE:				

QY	1	ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT	828639
Db	828698	ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT	
QY	61	ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT	828639
Db	828638	ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT	
QY	121	AGGCTCTATAGCAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT	828579
Db	828578	AGGCTCTATAGCAACAGGCTCGCATTAATTATTAGTCAGAGCTTTTTTTTATTTTTT	
QY	181	GGCGAGTCAGCAATAAAGAGTGGGAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC	240
Db	828518	GGCGAGTCAGCAATAAAGAGTGGGAGCTCAAGAAATAGCAGGTGCGGAAGCTAAGCC	828459
QY	241	TAAAGATCTAAGACCGATCTCTAGAGCGATGCGATCTTGGCTTCTGCGATGATGC	300
Db	828519	TAAAGATCTAAGACCGATCTCTAGAGCGATGCGATCTTGGCTTCTGCGATGATGC	

Query Match 99.4%; Score 2143.4; DB 15; Length 1230025;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2155; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY	1381	TAAAGCTGGGCAACAA	CAGGCA	TACTCAATGC	TTTAGGACAGATCGCTTCTGCTGCTG-	1439
Db	827318	TAAAGCTGGGCAACAA	CAGGCA	TACTCAATGC	TTTAGGACAGATCGCTTCTGCTGCTGT	827259
QY	1440	TTGTGAGGCGAGGATTC	CTCCCGCTGCAGCAAGTTCTAT	TAGGCTCATCTCTTAACAACGC		1499
Db	827258	TTGTGAGGCGAGGATTC	CTCCCGCTGCAGCAAGTTCTAT	TAGGCTCATCTCTTAACAACGC		827199
QY	1500	TTTACAAGACCTCAAAATCT	ACAGTTC	TAGATTATAAACA	CAGATATCAGCAGGTTATG	1559
Db	827198	TTTACAAGACCTCAAAATCT	ACAGTTC	TAGATTATAAACA	CAGATATCAGCAGGTTATG	827139
QY	1560	ATGCTTACAATCCATCAAT	GATGCCTATGTGTAGGCA	CGAAATGATCGCACTCGTGATG		1619
Db	827138	ATGCTTACAATCCATCAAT	GATGCCTATGTGTAGGCA	CGAAATGATCGCACTCGTGATG		827079
QY	1620	TGATAAACATGTAAGTAC	CCCCCGCTCTCA	CAGCATCCGTTCTCTAGAGCA	CGAACAGAAAG	1679
Db	827078	TGATAAACATGTAAGTAC	CCCCCGCTCTCTCA	CAGCATCCGTTCTCTAGAGCA	CGAACAGAAAG	827019
QY	1680	CTCAGAGCACGAAAAAAC	AGATCAAGCCCTCGCTAGG	TGATTCGCGCAATACGAA		1739
Db	827018	CTCAGAGCACGAAAAAAC	AGATCAAGCCCTCGCTAGG	TGATTCGCGCAATACGAA		826959
QY	1740	CTCTTGGAGATGTC	TATAGTCAAGTTTCGGCA	CTACAATCTGTAA	TGCAGATCATCCAGT	1799
Db	826958	CTCTTGGAGATGTC	TATAGTCAAGTTTCGGCA	CTACAATCTGTAA	TGCAGATCATCCAGT	826899
QY	1800	CGAATCCTCAAGCGAAT	AATGAGGAGATCAGACAA	AAAGCTTACATCGG	CAGTGACAAAGC	1859
Db	826898	CGAATCCTCAAGCGAAT	AATGAGGAGATCAGACAA	AAAGCTTACATCGG	CAGTGACAAAGC	826839
QY	1860	CTCCACAGTTTGGCTAT	CCCTTATGTGCAACTTTCTAAT	TGACTCTCTACACAGAA	GTTCATAG	1919
Db	826838	CTCCACAGTTTGGCTAT	CCCTTATGTGCAACTTTCTAAT	TGACTCTCTACACAGAA	GTTCATAG	826779
QY	1920	CTAAATAGAAAGTTTGTT	TGCTGCAAGGATCTAGGAC	AGCGCTGAAATAAAG	CACATTT	1979
Db	826778	CTAAATAGAAAGTTTGTT	TGCTGCAAGGATCTAGGAC	AGCGCTGAAATAAAG	CACATTT	826719
QY	1980	CCCTTTGAAACGAACTCC	CTTGTTTATTTACAGCAG	TGCTGCTCAATTCGGCTCTCTATATTT		2039
Db	826718	CCCTTTGAAACGAACTCC	CTTGTTTATTTACAGCAG	TGCTGCTCAATTCGGCTCTCTATATTT		826659
QY	2040	CTGCTTATCTCCAATA	CAACACCTTAAGTGTTCGTTT	GGAGAGATATTATATGTGCTTTGG		2099
Db	826658	CTGCTTATCTCCAATA	CAACACCTTAAGTGTTCGTTT	GGAGAGATATTATATGTGCTTTGG		826599
QY	2100	TAAAGCCCTTTGTTAGG	CGCTTACCAACACATAGAAC	CGATCTTCAATAATAAAGA		2156
Db	826598	TAAAGCCCTTTGTTAGG	CGCTTACCAACACATAGAAC	CGATCTTCAATAATAAAGA		826542

RESULT 3
US-09-841-132-385
; Sequence 385, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09841132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 385
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae

1121 GTTGGAGGCTCAAGCAACAGGAGTAGTATTGGTAGTATTGGTTTCCATGCTGTTA 1180
1021 GTTGGAGGCTCAAGCAACAGGAGTAGTATTGGTAGTATTGGTTTCCATGCTGTTA 1080
1181 GATGATGCTGAAAAATGAGACCGCTTCCATTTTGAATGCTGGGTTTGGTCCAGATGATTTAC 1240
1081 GATGATGCTGAAAAATGAGACCGCTTCCATTTTGAATGCTGGGTTTGGTCCAGATGATTTAC 1140
1241 ATGTTCAATACGAAAAATCTGATTTCTAAGCTGCCCAACAGAGCTCCGACGACAAGCT 1300
1141 ATGTTCAATACGAAAAATCTGATTTCTAAGCTGCCCAACAGAGCTCCGACGACAAGCT 1200
1301 AGAGCAGCAAGAACCGCTGGAGATGACAGTCTGTCGACGCTGGCAGATGCTCAGAAA 1360
1201 AGAGCAGCAAGAACCGCTGGAGATGACAGTCTGTCGACGCTGGCAGATGCTCAGAAA 1260
1361 GCTTTAGAAAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1420
1261 GCTTTAGAAAGCGGCTCTAGGTAAGCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGA 1320
1421 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTCTTCCGCTGCACGAACTTCTATA 1480
1321 CAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTCTTCCGCTGCACGAACTTCTATA 1380
1481 GGGTCATCTGTAACACAGCTTTACAAGACTCAAAATCTACAGGTTCTGATTTATAAACA 1540
1381 GGGTCATCTGTAACACAGCTTTACAAGACTCAAAATCTACAGGTTCTGATTTATAAACA 1440
1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCCCTATGGTAGGCGACA 1600
1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCCCTATGGTAGGCGACA 1500
1601 AATGATGCGACTCGTCTGATGTAACAAATGTAAGTACCCGCTCTCACAGATCCGTT 1660
1501 AATGATGCGACTCGTCTGATGTAACAAATGTAAGTACCCGCTCTCACAGATCCGTT 1560
1661 CCTAGAGCAGCAACAGAGCTCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1720
1561 CCTAGAGCAGCAACAGAGCTCGAGGACCAAGAAAAACAGATCAAGCCCTCGCTAGGGTG 1620
1721 ATTCTGGCAATAGCAGAACTTTGAGATGTTCTATAGTCAAGTTTTCGGCCTACAACT 1780
1621 ATTCTGGCAATAGCAGAACTTTGAGATGTTCTATAGTCAAGTTTTCGGCCTACAACT 1680
1781 GTAATCAGATCATCCAGTCGAATCTCAAGCGAAATAGAGGAGATCAGAAAAAGCTT 1840
1681 GTAATCAGATCATCCAGTCGAATCTCAAGCGAAATAGAGGAGATCAGAAAAAGCTT 1740
1841 ACATCGGCAAGTGAACAAAGCTCCACAGTTTGGCTATCCCTTATGTGCAACTTTCTAATGAC 1900
1741 ACATCGGCAAGTGAACAAAGCTCCACAGTTTGGCTATCCCTTATGTGCAACTTTCTAATGAC 1800
1901 TCTACAGAGAGTTCTAGCTAAATTAGAAGTTTCTTGGTGAAGAGTCTAGGACGCA 1960
1801 TCTACAGAGAGTTCTAGCTAAATTAGAAGTTTCTTGGTGAAGAGTCTAGGACGCA 1860
1961 GCTGAATAAAGCACTTTCCTTTGAAACGAACCTCTTGTATTATCAGCAGAGTCTGGTTC 2020
1861 GCTGAATAAAGCACTTTCCTTTGAAACGAACCTCTTGTATTATCAGCAGAGTCTGGTTC 1920
2021 AATATCGGCTCTCTATATTCTGTTATCTGTTATCTCCATAA 2056
1921 AATATCGGCTCTCTATATTCTGTTATCTCTCCATAA 1956

RESULT 4

US-09-841-260-63
; Sequence 63, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter

APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; FILE REFERENCE: 21/121.515
; CURRENT APPLICATION NUMBER: US/09/841.260
; CURRENT FILING DATE: 2004-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-63

Query Match 90.7%; Score 1956; DB 10; Length 1956;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 160
DB 1 ATGTTAATCTTATTTGGTCCAGGTCCTATAGACGAAACAGAACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATATAGAGTCCGGAAGCTCAAGAATA 220
DB 61 CTTTCTGCTCAAGGATTTGGAGCGAGTGCAGCAATATAGAGTCCGGAAGCTCAAGAATA 120
QY 221 GCAGTGGCGAAGCTTAAGCTAAAGAACTTAAGACCGGATTTCTGTAGAGCGATGGAGCATC 280
DB 121 GCAGTGGCGAAGCTTAAGCTAAAGAACTTAAGACCGGATTTCTGTAGAGCGATGGAGCATC 180
QY 281 TTGCGTTCTGCAAGTGAATGCTCTCATGAGTCTGGCAGATATAGCTGGGTATTTGCTTTAGT 340
DB 181 TTGCGTTCTGCAAGTGAATGCTCTCATGAGTCTGGCAGATATAGCTGGGTATTTGCTTTAGT 240
QY 341 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTTGACTCAACGACGACCGACCT 400
DB 241 AACAGCTGCTCTTCTACTAGCAGATCTGCAGACGTTGACTCAACGACGACCGACCT 300
QY 401 AGCGTCTCTCCACCCACCGTTTGATGATTATAGACTCAAGCGCAACAGCTTACGATAT 460
DB 301 AGCGTCTCTCCACCCACCGTTTGATGATTATAGACTCAAGCGCAACAGCTTACGATAT 360
QY 461 ATCTTTAATCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
DB 361 ATCTTTAATCTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
QY 521 GTCACATAATATAAGATACAGCGCTACTGATAGGAAACCGCAATCGCTCGGAGTGG 580
DB 421 GTCACATAATATAAGATACAGCGCTACTGATAGGAAACCGCAATCGCTCGGAGTGG 480
QY 581 GAAACTAAGAAATGCCGATGCACTTAAAGTTGGCGGCAAAATTAAGAAATAGCGAAATAT 640
DB 481 GAAACTAAGAAATGCCGATGCACTTAAAGTTGGCGGCAAAATTAAGAAATAGCGAAATAT 540
QY 641 GCTTCGGATAACCAAGGATTTCTGACTTTTGGTAACTGACTTCTCCCTCCAGCTCTTA 700
DB 541 GCTTCGGATAACCAAGGATTTCTGACTTTTGGTAACTGACTTCTCCCTCCAGCTCTTA 600
QY 701 CAGGCTGCTCTTCTCCATCTCTAGCAACAAATAAAGACGCTGAGCTTCTTAAAGAG 760
DB 601 CAGGCTGCTCTTCTCCATCTCTAGCAACAAATAAAGACGCTGAGCTTCTTAAAGAG 560
QY 761 ATGCAAGATAACCCAGTAGTCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTAGT 820
DB 661 ATGCAAGATAACCCAGTAGTCCAGGAAACCGCTGCAATTTGCTCAATCTTTAGTTAGT 720
QY 821 CAGACAGATGCTTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCAAT 880
DB 721 CAGACAGATGCTTACAGCGACACAGATAGAGAAAGATGGAATGCGATTAGGATGCAAT 780
QY 881 TTTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAAATGCTAAATCTAATAAAGATTAAGC 940
DB 781 TTTTGAGGACAGAACGCTAGTGGAGCTGTAGAAAAATGCTAAATCTAATAAAGATTAAGC 840

941 AACATAGATTACGCTAAAGCAGCAATCGCTACTGCTAAGACAAATAGCTGAAGCTCAG 1000
Db
841 AACATAGATTACGCTAAAGCAGCAATCGCTACTGCTAAGACAAATAGCTGAAGCTCAG 900
QY
1001 AAAAGTCTCCCGACTCTCAATCTTCAAGAGCGGAAACAAATGGTAATAACAGCTGAG 1060
Db
901 AAAAGTCTCCCGACTCTCAATCTTCAAGAGCGGAAACAAATGGTAATAACAGCTGAG 960
QY
1061 AAAGATCTTAAATAATCAAACTGAGATGTTCTGATGTTCCAAATCCAGGAATACA 1120
Db
961 AAAGATCTTAAATAATCAAACTGAGATGTTCTGATGTTCCAAATCCAGGAATACA 1020
QY
1121 GTTGAGGCTCCAAGCAACAGGAAGTAGTATTGGTAGTATTGGTGTTCCTGCTGTTA 1180
Db
1021 GTTGAGGCTCCAAGCAACAGGAAGTAGTATTGGTAGTATTGGTGTTCCTGCTGTTA 1080
QY
1181 GATGATGCTGAAAATGAGACCGCTTCCATTTTGAATGCTGGGTTTGGTCAAGATGATC 1240
Db
1081 GATGATGCTGAAAATGAGACCGCTTCCATTTTGAATGCTGGGTTTGGTCAAGATGATC 1140
QY
1241 ATGTTCAATACGGAATCCTGATTTCTCAAGCTGCCCAAGAGCTGCAGCAAGCT 1300
Db
1141 ATGTTCAATACGGAATCCTGATTTCTCAAGCTGCCCAAGAGCTGCAGCAAGCT 1200
QY
1301 AGAGCAGCAAGAGCGCTGGAGATGACAGTGTCTGCGAGCGCTGGAGATGCTCAGAAA 1360
Db
1201 AGAGCAGCAAGAGCGCTGGAGATGACAGTGTCTGCGAGCGCTGGAGATGCTCAGAAA 1260
QY
1361 GCTTTAGAGCGGCTTAGGTAAAGTGGGCAACAGAGGATGATCTCAATGCTTAGGA 1420
Db
1261 GCTTTAGAGCGGCTTAGGTAAAGTGGGCAACAGAGGATGATCTCAATGCTTAGGA 1320
QY
1421 CAGATCGCTTCTGCTGCTGTGAGCGCAGAGTTCCCTCCGCTGCAGCAAGTTCTATA 1480
Db
1321 CAGATCGCTTCTGCTGCTGTGAGCGCAGAGTTCCCTCCGCTGCAGCAAGTTCTATA 1380
QY
1481 GGGTCATCTGTAACAGCTTTAAGACCTCAAAATCTACAGGTTCTGATTAATAACA 1540
Db
1381 GGGTCATCTGTAACAGCTTTAAGACCTCAAAATCTACAGGTTCTGATTAATAACA 1440
QY
1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCCCTATGGTAGGCA 1600
Db
1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCCCTATGGTAGGCA 1500
QY
1601 AATGATGCGACTCGTGATGATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTT 1660
Db
1501 AATGATGCGACTCGTGATGATAAACAATGTAAGTACCCCGCTCTCACAGATCCGTT 1560
QY
1661 CTTAGAGCAGAACAGAGCTCGAGACCCAGAAAAACAGATCAAGCCCTCGTAGGGTG 1720
Db
1561 CTTAGAGCAGAACAGAGCTCGAGACCCAGAAAAACAGATCAAGCCCTCGTAGGGTG 1620
QY
1721 ATTCTGCGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTCGGCACTCAATCT 1780
Db
1621 ATTCTGCGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTCGGCACTCAATCT 1680
QY
1781 GTAATGCGAGATCATCCAGTCGAATCTCAAGCGAATTAATGAGGAGATCAGACAAAGCTT 1840
Db
1681 GTAATGCGAGATCATCCAGTCGAATCTCAAGCGAATTAATGAGGAGATCAGACAAAGCTT 1740
QY
1841 ACATGGCAGTGCAGAAAGCTCAGAGTTTGGCTATCTTATGTCGCACTTCTAATGAC 1900
Db
1741 ACATGGCAGTGCAGAAAGCTCAGAGTTTGGCTATCTTATGTCGCACTTCTAATGAC 1800
QY
1901 TCTACACAGAGTTCTATAGTAAATAGAAAGTTTGTGCTGAAGGATCTAGGACAGCA 1960
Db
1801 TCTACACAGAGTTCTATAGTAAATAGAAAGTTTGTGCTGAAGGATCTAGGACAGCA 1860
QY
1961 GCTGAAATAAAGCACTTTCTTTGAAACGAACTCCTTGTATTAACAGAGGCTGCTGCTC 2020
Db
1861 GCTGAAATAAAGCACTTTCTTTGAAACGAACTCCTTGTATTAACAGAGGCTGCTGCTC 1920
QY
2021 AATATCGGCTCTCTATATTCTGGTTATCTCCAAATA 2056

Db 1921 AATATCGGCTCTCTATATTCTGGTTATCTCCAAATA 1956

RESULT 5

US-10-007-693-63
; Sequence 63, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210471-545C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-63

Query Match 90.7%; Score 1956; DB 13; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATGGTTAATCTATTGGTCCAGGTCCTATAGACGAAACAGACGACACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCTATTGGTCCAGGTCCTATAGACGAAACAGACGACACCTCCCGCAGAT 60
QY 161 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATAAAGAGTGCAGAAAGCTCAAAAGATA 220
Db 61 CTTTCTGCTCAAGGATTGGAGCGAGTGCAGCAATAAAGAGTGCAGAAAGCTCAAAAGATA 120
QY 221 GCAGTGGGAAAGCTTAAGCTTAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGATC 280
Db 121 GCAGTGGGAAAGCTTAAGCTTAAAGATCTAAGACCGATTCTGTAGAGCGATGGAGATC 180
QY 281 TTGGTCTCGAGTGAATGCTCTCATGAGTCTGCAGATAAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGGTCTCGAGTGAATGCTCTCATGAGTCTGCAGATAAGCTGGGTATTGCTTCTAGT 240
QY 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACAGACGACGACCGACT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACAGACGACGACCGACT 300
QY 401 AGGCTCTCTCCACCCAGCTTTGATGATTAAGACTCAAGCGCAAAACAGCTTACGATACT 460
Db 301 AGGCTCTCTCCACCCAGCTTTGATGATTAAGACTCAAGCGCAAAACAGCTTACGATACT 360
QY 461 ATCTTTACTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 520
Db 361 ATCTTTACTCAACATCACTAGCTGACATACAGGCTGCTTTGGTGAGCTCCAGGATGCT 420
QY 521 GTCACATAATAAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 580
Db 421 GTCACATAATAAAGGATACAGCGGCTACTGATGAGGAAACCGCAATCGCTGCGAGTGG 480
QY 581 GAACTAAGATGCGGATGCGAGTTAAAGTTGCGCGCAAAATACAGAAATGAGGAAATAT 640
Db 481 GAACTAAGATGCGGATGCGAGTTAAAGTTGCGCGCAAAATACAGAAATGAGGAAATAT 540
QY 641 GCTTCGGATAACCAAGCGATTCTTGACTTTTAGTAACTGACTCTCTCCGACTCTTA 700
Db 541 GCTTCGGATAACCAAGCGATTCTTGACTTTTAGTAACTGACTCTCTCCGACTCTTA 600
QY 701 CAGGCTGCTCTCTCCAAATCTGTAGCAAAACAAATACAAAGCAGCTGAGCTTCTTAAAGAG 760
Db 601 CAGGCTGCTCTCTCCAAATCTGTAGCAAAACAAATACAAAGCAGCTGAGCTTCTTAAAGAG 660
QY 761 ATGCMAGTAACCCAGTAGTCCCGAGGAAACCGCTGCAATGCTCAATCTTTAGTTCAT 820

Db 661 ATGCAAGATAACCCAGTAGTCCAGGGAAACCGCTGCAATTGCTCAATCTTTAGTTGAT 720
Qy 821 CAGACAGATGCTACAGCAGACAGATAGAGAAAGATGAAATGCGATTAGGATGCATAT 880
Db 721 CAGACAGATGCTACAGCAGACAGATAGAGAAAGATGAAATGCGATTAGGATGCATAT 780
Qy 881 TTTGAGGACAGAACCGCTAGTGGAGCTGTAGAAATGCTAAATCTAAATCAAGTATAAGC 940
Db 781 TTTGAGGACAGAACCGCTAGTGGAGCTGTAGAAATGCTAAATCTAAATCAAGTATAAGC 840
Qy 941 AACATAGATTACGCTTAAGAGAGCAATCGCTACTGTCTAAGACACAAATAGCTGAAGCTCAG 1000
Db 841 AACATAGATTACGCTTAAGAGAGCAATCGCTACTGTCTAAGAGACACAAATAGCTGAAGCTCAG 900
Qy 1001 AAAAGATTCCCGGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGCTGAG 1060
Db 901 AAAAGATTCCCGGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGCTGAG 960
Qy 1061 AAAAGATTCTTAAATAATCAAACTCGAGATGTTCTGATGTTCCAAATCCAGGAACCTACA 1120
Db 961 AAAAGATTCTTAAATAATCAAACTCGAGATGTTCTGATGTTCCAAATCCAGGAACCTACA 1020
Qy 1121 GTTGAGGCTCAAGCAGACAGGAAGTAGTAGTGTAGTAGTATGCGTTTCCATGCTGTTA 1180
Db 1021 GTTGAGGCTCAAGCAGACAGGAAGTAGTAGTGTAGTAGTATGCGTTTCCATGCTGTTA 1080
Qy 1181 GATGATGCTGAAATGAGACCGCTTCCATTTTGTGTTCTGTTGTTTGGTTCGTCAGATGTTCC 1240
Db 1081 GATGATGCTGAAATGAGACCGCTTCCATTTTGTGTTCTGTTGTTTGGTTCGTCAGATGTTCC 1140
Qy 1241 ATGTTCAATACGGAATCTCTGATTCTCAAGTGCCCAACAGGAGCTCGCAGCAAGCT 1300
Db 1141 ATGTTCAATACGGAATCTCTGATTCTCAAGTGCCCAACAGGAGCTCGCAGCAAGCT 1200
Qy 1301 AGAGCAGGAAAGCGCTGAGATGACAGTGTCTGTCAGAGCTGCGAGTGTCTCAGAAA 1360
Db 1201 AGAGCAGGAAAGCGCTGAGATGACAGTGTCTGTCAGAGCTGCGAGTGTCTCAGAAA 1260
Qy 1361 GCTTTAGAGCGCTCTAGGTAAAGCTGGCAACAAAGGAGCTACTCAATGCTTTAGGA 1420
Db 1261 GCTTTAGAGCGCTCTAGGTAAAGCTGGCAACAAAGGAGCTACTCAATGCTTTAGGA 1320
Qy 1421 CAGATCGCTTCTGCTGTGTGTGTAGCGCAGAGTTCCTCCGCTGCGAGCAAGTTCTATA 1480
Db 1321 CAGATCGCTTCTGCTGTGTGTGTAGCGCAGAGTTCCTCCGCTGCGAGCAAGTTCTATA 1380
Qy 1481 GGGTCATCTGTAAACAGCTTTACAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1540
Db 1381 GGGTCATCTGTAAACAGCTTTACAGACCTCAAAATCTACAGGTTCTGATTATAAACA 1440
Qy 1541 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTGATGAGGCAAGA 1600
Db 1441 CAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTGATGAGGCAAGA 1500
Qy 1601 AATGATGCACTGCTGTGTGTGTAAACATGTAAGTACCCCGCTCTCAGAGATCCGTT 1660
Db 1501 AATGATGCACTGCTGTGTGTGTAAACATGTAAGTACCCCGCTCTCAGAGATCCGTT 1560
Qy 1661 CCTAGGACAGCAACAGAGCTCGAGGACAGAAACAGATCAAGCCCTCGCTAGGGTG 1720
Db 1561 CCTAGGACAGCAACAGAGCTCGAGGACAGAAACAGATCAAGCCCTCGCTAGGGTG 1620
Qy 1721 ATTCTTGGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTGGCACTCAAACT 1780
Db 1621 ATTCTTGGCAATAGCAGAACTCTTGGAGATGCTATAGTCAAGTTTGGCACTCAAACT 1680
Qy 1781 GTAATGCAATATCCAGTGCATCTTCAAGCAATATAGGAGATCAGACAAAGCTT 1840
Db 1681 GTAATGCAATATCCAGTGCATCTTCAAGCAATATAGGAGATCAGACAAAGCTT 1740
Qy 1841 ACATCGGAGTGCAGAAAGCTTCCCAAGCTTGGCTATGCTATGCTCACTTCTAATGAC 1900
Db 1741 ACATCGGAGTGCAGAAAGCTTCCCAAGCTTGGCTATGCTATGCTCACTTCTAATGAC 1800

Qy 1901 TCTACACAGAGTTTCATAGTAATAGAAAGTTTGTTCGTAAGGATCTAGGACAGCA 1960
Db 1801 TCTACACAGAGTTTCATAGTAATAGAAAGTTTGTTCGTAAGGATCTAGGACAGCA 1860
Qy 1961 GCTGAAATAAAAGCACCTTCTTTGAAACGAACTCTTGTTTTATTCAGCAGGTCGTC 2020
Db 1861 GCTGAAATAAAAGCACCTTCTTTGAAACGAACTCTTGTTTTATTCAGCAGGTCGTC 1920
Qy 2021 AATATCGGCTCTTATATTTCTGTTTCTTCTCAATAA 2056
Db 1921 AATATCGGCTCTTATATTTCTGTTTCTTCTCAATAA 1956

RESULT 6

US-10-312-273-14
; Sequence 14, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035MO
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-28
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017989.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 14
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-14

Query Match 90.7%; Score 1956; DB 15; Length 1956;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1956; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 ATGGTTAATCCTATTGGTCCAGGTCCTATAGCAAGAAACAGAAACGACACCTCCCGCAGAT 160
Db 1 ATGGTTAATCCTATTGGTCCAGGTCCTATAGCAAGAAACAGAAACGACACCTCCCGCAGAT 60
Qy 161 CTTTCTGTCTCAAGGATTGGAGCGAGTGCAGCAATAAGAGTGCAGAGCTCAAGAATA 220
Db 61 CTTTCTGTCTCAAGGATTGGAGCGAGTGCAGCAATAAGAGTGCAGAGCTCAAGAATA 120
Qy 221 GCAGTGGGAAGCTAAGCTTAAAGATCTAAGACCGATCTGTAGAGCGATGAGCATC 280
Db 121 GCAGTGGGAAGCTAAGCTTAAAGATCTAAGACCGATCTGTAGAGCGATGAGCATC 180
Qy 281 TTGGTCTTCAGAGTAATGCTCTCATGAGTCTGCAGATAGCTGGGTATTGCTTCTAGT 340
Db 181 TTGGTCTTCAGAGTAATGCTCTCATGAGTCTGCAGATAGCTGGGTATTGCTTCTAGT 240
Qy 341 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGAGACCGCACT 400
Db 241 AACAGCTCGTCTTCTACTAGCAGATCTGCAGACGTGGACTCAACGACGAGACCGCACT 300
Qy 401 AGGCTCTCCACCCACGTTTGTGATTAAGACTCAGCGCAACAGCTTACGATACT 460

off

Db 1 ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 60
QY 365 TCTGACAGAGTGGACTCAACGACGACGACCGACCTAGCGCTCTCCACCCACGCTTTGAT 424
Db 61 TCTGACAGAGTGGACTCAACGACGACGACCGACCGACCTAGCGCTCTCCACCCACGCTTTGAT 120
QY 425 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTATCTTACCTCAACATCACTAGCT 484
Db 121 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTATCTTACCTCAACATCACTAGCT 180
QY 485 GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACCTAATAAAGGATACAGCG 544
Db 181 GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACCTAATAAAGGATACAGCG 240
QY 545 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGAACTTAAGAAATGCCGATCGATT 604
Db 241 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGAACTTAAGAAATGCCGATCGATT 300
QY 605 AAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGGATTCCT 664
Db 301 AAAAGTTGGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATAACCAAGCGGATTCCT 360
QY 665 GACTCTTTAGGTAAGTACTCTTCTTGGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA 724
Db 361 GACTCTTTAGGTAAGTACTCTTCTTGGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA 420
QY 725 GCAAACTAATAACAGCAGCTGAGCTCTTAAAGAGATGCAAGATAACCCAGTACTGCCA 784
Db 421 GCAAACTAATAACAGCAGCTGAGCTCTTAAAGAGATGCAAGATAACCCAGTACTGCCA 480
QY 785 GGGAAAAAGCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 844
Db 481 GGGAAAAAGCCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
QY 845 ATAGAGAAAGATGGAATGCGAATAGGATGATGATATTTTGGAGCAGAAACGCTAGTGA 904
Db 541 ATAGAGAAAGATGGAATGCGAATAGGATGATGATATTTTGGAGCAGAAACGCTAGTGA 600
QY 905 GCTGTAGAAAATGCTAAATCTAATACAGTAAAGCAATAGATTCAGTAAAGCAGCA 964
Db 601 GCTGTAGAAAATGCTAAATCTAATACAGTAAAGCAATAGATTCAGTAAAGCAGCA 660
QY 965 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCGCCGACTCTCCAAT 1024
Db 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCGCCGACTCTCCAAT 720
QY 1025 CTTCAAGAGCGGAACAAATGTTATACAGGCTGAGAAAGATCTTAAATAATCAAACT 1084
Db 721 CTTCAAGAGCGGAACAAATGTTATACAGGCTGAGAAAGATCTTAAATAATCAAACT 780
QY 1085 GCAGATGGTTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGCA 1144
Db 781 GCAGATGGTTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGCA 840
QY 1145 AGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1204
Db 841 AGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 1205 TCCATTTTATGCTGCTGGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
Db 901 TCCATTTTATGCTGCTGGGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 1265 TCTCAAGCTGCCCAACAGGAGCTCGCAGACACAGCTAGACGAGCAAGCCGCTGGAGAT 1324
Db 961 TCTCAAGCTGCCCAACAGGAGCTCGCAGACACAGCTAGACGAGCAAGCCGCTGGAGAT 1020
QY 1325 GACAGTGTCTGCGACGCTGCGAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA 1384
Db 1021 GACAGTGTCTGCGACGCTGCGAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA 1080
QY 1385 GCTGGGCAACACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGG 1444

Db 1081 GCTGGCAACACAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGGT 1140
QY 1445 AGCGCAGAGTCTCTCCGCTGCGAGCAAGTCTTATAGGTCATCTGTAAACACGCTTTAC 1504
Db 1141 AGCGCAGAGTCTCTCCGCTGCGAGCAAGTCTTATAGGTCATCTGTAAACACGCTTTAC 1200
QY 1505 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTATGATGCT 1564
Db 1201 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGCAGGTTATGATGCT 1260
QY 1565 TACAATTCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGACTCTGATGATGATA 1624
Db 1261 TACAATTCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGACTCTGATGATGATA 1320
QY 1625 AACAAATGTAAGTACCCCGCTCTCACAGCTCCGTTCTAGAGCACGAACAGAAAGCTCGA 1684
Db 1321 AACAAATGTAAGTACCCCGCTCTCACAGCTCCGTTCTAGAGCACGAACAGAAAGCTCGA 1380
QY 1685 GGACCAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCTT 1744
Db 1381 GGACCAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCTT 1440
QY 1745 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCGAGATCATCCAGTCAAT 1804
Db 1441 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCGAGATCATCCAGTCAAT 1500
QY 1805 CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCGAGTGACAAAAGCTCCA 1864
Db 1501 CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCGAGTGACAAAAGCTCCA 1560
QY 1865 CAGTTTGGCTATCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTCCATAGCTAAA 1924
Db 1561 CAGTTTGGCTATCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTCCATAGCTAAA 1620
QY 1925 TTAGAAAGTTTGTGCTGAGGATCTAGGACAGCTGAAATGAAAGCACTTTCCCTTT 1984
Db 1621 TTAGAAAGTTTGTGCTGAGGATCTAGGACAGCTGAAATGAAAGCACTTTCCCTTT 1680
QY 1985 GAAACGAACTCCTTGTATTTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCTGGT 2044
Db 1681 GAAACGAACTCCTTGTATTTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCTGGT 1740
QY 2045 TATCTCCAAATAACACACTAAGTGTGCTTTGGAGAGATTTATGTTGCTTTGTTAGG 2104
Db 1741 TATCTCCAAATAACACACTAAGTGTGCTTTGGAGAGATTTATGTTGCTTTGTTAGG 1800
QY 2105 CTTTGTGTGGGCTTACCAACACACTAGACGATCTTCAATAATAAAGA 2156
Db 1801 CTTTGTGTGGGCTTACCAACACACTAGACGATCTTCAATAATAAAGA 1852

RESULT 8

US-09-564-479-7
; Sequence 7, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COHEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: US/09/564,479
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2238

```
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (766)..(2235)
US-09-564-479-7

Query Match      68.0%; Score 1465.8; DB 10; Length 2238;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1478; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAAGAGCTTTTATTTTATTTT 60
Db 666 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAAGAGCTTTTATTTTATTTT 725

QY 61 ATATATAACTTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 120
Db 726 ATATATAACTTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 785

QY 121 AGGTCTTATAGACGAAACAGAACGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 180
Db 786 AGGTCTTATAGACGAAACAGAACGACACCTCCGCGAGATCTTTCTGCTCAAGGATTGGA 845

QY 181 GCGAGTGCAGCAATAATAGATGCGGAAGCTCAAGAAATAGCAGTGCAGGAAGCTTAAGCC 240
Db 846 GCGAGTGCAGCAATAATAGATGCGGAAGCTCAAGAAATAGCAGTGCAGGAAGCTTAAGCC 905

QY 241 TAAAGATCTAAGACCGATCTGTAGAGCGATGGAGCATCTTGGTCTCAGTGAATGC 300
Db 906 TAAAGATCTAAGACCGATCTGTAGAGCGATGGAGCATCTTGGTCTCAGTGAATGC 965

QY 301 TCTCATAGTCTGCGAGATAAGCTGGGTATTGCTTTAGTAAACAGCTCGTCTTCTACTAG 360
Db 966 TCTCATAGTCTGCGAGATAAGCTGGGTATTGCTTTAGTAAACAGCTCGTCTTCTACTAG 1025

QY 361 CAGATCTGCAGACGTGAGTCAACAGACAGACGACGACCTAGCCTCTCTCACCCAGCTT 420
Db 1026 CAGATCTGCAGACGTGAGTCAACAGACAGACGACGACCTAGCCTCTCTCACCCAGCTT 1085

QY 421 TGATGATTATAGACTCAAGCGCAACAGCTTACGATACTATCTTTACTCTCAACATCACT 480
Db 1086 TGATGATTATAGACTCAAGCGCAACAGCTTACGATACTATCTTTACTCTCAACATCACT 1145

QY 481 AGCTGACATACAGGCTGCTTTGGTGAGCTCCTCAGGATGCTCTCACTAATAAAGGATAC 540
Db 1146 AGCTGACATACAGGCTGCTTTGGTGAGCTCCTCAGGATGCTCTCACTAATAAAGGATAC 1205

QY 541 AGCGGCTACTGATGAGGAAACGGCAATCGCTCGGAGTGGGAACATAGAAATGCCGATGC 600
Db 1206 AGCGGCTACTGATGAGGAAACGGCAATCGCTCGGAGTGGGAACATAGAAATGCCGATGC 1265

QY 601 AGTTAAAGTTGCGGCGCAATTTACAGATTAGCGAAATATGCTTCGGATAACCAAGCGAT 660
Db 1266 AGTTAAAGTTGCGGCGCAATTTACAGATTAGCGAAATATGCTTCGGATAACCAAGCGAT 1325

QY 661 TCTTGACTCTTTAGTAAACTGACTCTTCCGACTCTTCAAGCTCTTCAAGCTCTTCTCCATC 720
Db 1326 TCTTGACTCTTTAGTAAACTGACTCTTCCGACTCTTCAAGCTCTTCTCCATC 1385

QY 721 TGTACCAACATATACAGAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAT 780
Db 1386 TGTACCAACATATACAGAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAT 1445

QY 781 CCCAGGAAACCGCTCGCAATGCTCAATCTTTAGTGTATGATCAGACAGATGCTACAGCGAC 840
Db 1446 CCCAGGAAACCGCTCGCAATGCTCAATCTTTAGTGTATGATCAGACAGATGCTACAGCGAC 1505

QY 841 ACAGATAGAGAAGATGGAATTCGATTAGGATGCATATTTTGCAGACAGAACGCTAG 900
Db 1506 ACAGATAGAGAAGATGGAATTCGATTAGGATGCATATTTTGCAGACAGAACGCTAG 1565

QY 901 TGGAGCTCTAGAAATGCTTAATCTTAATAACAGTATAGCAACATAGATTAGCTTAAGCC 960
Db 901 TGGAGCTCTAGAAATGCTTAATCTTAATAACAGTATAGCAACATAGATTAGCTTAAGCC 960
```

```
Db 1566 TGGAGCTCTAGAAATGCTTAATCTTAATAACAGTATAGCAACATAGATTAGCTTAAGCC 1625
QY 961 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGGACTCTCC 1020
Db 1626 AGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGGACTCTCC 1685
QY 1021 AATTCTTCAAGAAGCGGAACAAATGTAATACAGGCTGAGAAAAGATCTTAAAAATATCAA 1080
Db 1686 AATTCTTCAAGAAGCGGAACAAATGTAATACAGGCTGAGAAAAGATCTTAAAAATATCAA 1745
QY 1081 ACCTGCAGATGTTCTGATGTTCCAAATCCAGNACTACAGTTGGAGGCTCCAGACCAACA 1140
Db 1746 ACCTGCAGATGTTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGACCAACA 1805
QY 1141 AGGAAGTAGTATTGGTATGTTTCCATGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 1200
Db 1806 AGGAAGTAGTATTGGTATGTTTCCATGTTTCCATGCTGTAGATGATGCTGAAAAATGAGAC 1865
QY 1201 CGCTTCCATTTGATGTTCTGGTTTCCTCAGATGATTCAATGTTCAATACGGAATATCC 1260
Db 1866 CGCTTCCATTTGATGTTCTGGTTTCCTCAGATGATTCAATGTTCAATACGGAATATCC 1925
QY 1261 TGATTTCAAGCTGCCCCAACAGGAGCTCGCAGCAACAGCTAGACGAGGAAGCGGCTGG 1320
Db 1926 TGATTTCAAGCTGCCCCAACAGGAGCTCGCAGCAACAGCTAGACGAGGAAGCGGCTGG 1985
QY 1321 AGATGACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTTAGG 1380
Db 1986 AGATGACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTTAGG 2045
QY 1381 TAAAGCTGGGCAACAAACAGGGCATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGT 1440
Db 2046 TAAAGCTGGGCAACAAACAGGGCATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGT 2105
QY 1441 TGTGAGCGCAGGAGTCTCTCCGCTGCAGCAAGTTCTATAG 1481
Db 2106 TGTGAGCGCAGGAG-TACTCCGCTGCAGCAAGTTCTATGG 2145

RESULT 9
US-09-564-479-5
; Sequence 5, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OOMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1456
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1456)
US-09-564-479-5

Query Match      67.5%; Score 1455; DB 10; Length 1456;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATATTAGTGAAGAGCTTTTATTTTATTTT 60
```

```
Db      1  ATAAATCTTTAAACACAGGCTCGATTAATTAATAGTGAGAGCTTTTATTTATTTT 60
QY      61  ATATAAATACTAAAGATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 120
Db      61  ATATAAATACTAAAGATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 120
QY     121  AGTCTCTATAGACGAAACAGACGACACCTCCCGAGATCTTTCTGCTCAAGGATTGGA 180
Db     121  AGTCTCTATAGACGAAACAGACGACACCTCCCGAGATCTTTCTGCTCAAGGATTGGA 180
QY     181  GCGAGTGCAGCAATAAAGAGTCCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 240
Db     181  GCGAGTGCAGCAATAAAGAGTCCGGAAGCTCAAGAATAGCAGGTGCGGAAGCTAAGCC 240
QY     241  TAAAGATCTAAGACCGATTCTGTAGAGGATGGAGCATCTTGGCTTCTGAGTGAATGC 300
Db     241  TAAAGATCTAAGACCGATTCTGTAGAGGATGGAGCATCTTGGCTTCTGAGTGAATGC 300
QY     301  TCTCATGAGTCTGGCAGATAAGCTGGGTATTCCTTAGTAACAGCTGCTCTTCTACTAG 360
Db     301  TCTCATGAGTCTGGCAGATAAGCTGGGTATTCCTTAGTAACAGCTGCTCTTCTACTAG 360
QY     361  CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACTAGCCCTCCACCACCGTT 420
Db     361  CAGATCTGCAGACGTGGACTCAACGACAGCGACCGCACTAGCCCTCCACCACCGTT 420
QY     421  TGATGATTAATAGACTCAAGCGCAACACAGCTTACGATACATCTTTTACCTCAACATCACT 480
Db     421  TGATGATTAATAGACTCAAGCGCAACACAGCTTACGATACATCTTTTACCTCAACATCACT 480
QY     481  AGCTGACATACAGGCTGCTTTGGTAGGCTCCAGGATGCTGCTCACTAATATAAGGATAC 540
Db     481  AGCTGACATACAGGCTGCTTTGGTAGGCTCCAGGATGCTGCTCACTAATATAAGGATAC 540
QY     541  AGCGGCTACTGATAGGAAACCGCAATCTGCTGGAGTGGGAACTAAGAAATGCGGATGC 600
Db     541  AGCGGCTACTGATAGGAAACCGCAATCTGCTGGAGTGGGAACTAAGAAATGCGGATGC 600
QY     601  AGTTAAAGTTGCGCGCAATACAGAAATAGCGAAATGCTCGGATTAACCAAGCCGAT 660
Db     601  AGTTAAAGTTGCGCGCAATACAGAAATAGCGAAATGCTCGGATTAACCAAGCCGAT 660
QY     661  TCTTGACTCTTTAGSTAAACTGACTCTCTTCCAGCTTACAGGCTGCTCTTCTCCAATC 720
Db     661  TCTTGACTCTTTAGSTAAACTGACTCTCTTCCAGCTTACAGGCTGCTCTTCTCCAATC 720
QY     721  TGTAGCAAACTAACAAGACGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
Db     721  TGTAGCAAACTAACAAGACGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGT 780
QY     781  CCCAGGAAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
Db     781  CCCAGGAAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGAC 840
QY     841  ACAGATAGAGAAGATGGAATGCGATAGGATGCAATTTTGCAGGACAGAAACGCTAG 900
Db     841  ACAGATAGAGAAGATGGAATGCGATAGGATGCAATTTTGCAGGACAGAAACGCTAG 900
QY     901  TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACGTTAAGC 960
Db     901  TGGAGCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACGTTAAGC 960
QY     961  AGCAATCGCTACTGTCAAGACACAAATAGCTGAAGCTCAAGAAAGTCTCCCGACTCTCC 1020
Db     961  AGCAATCGCTACTGTCAAGACACAAATAGCTGAAGCTCAAGAAAGTCTCCCGACTCTCC 1020
QY    1021  AATTCTTCAAGAGCGGAAACAAATGTTAATACAGGCTGAGAAAGATCTTTAAATAATCAA 1080
Db    1021  AATTCTTCAAGAGCGGAAACAAATGTTAATACAGGCTGAGAAAGATCTTTAAATAATCAA 1080
QY    1081  AACTGCAGATGGTCTTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACA 1140
```

```
Db     1081  ACCTGCAGATGGTCTGATGTTCCAAATCCAGGAACACAGTTGGAGGCTCCAAAGCAACA 1140
QY    1141  AGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
Db    1141  AGGAAGTAGTATTGGTAGTATTCTGTTTCCATGCTGTAGATGATGCTGAAAATGAGAC 1200
QY    1201  CGCTTCCATTTTGTATGCTGGGTTTCGTACATGATTCACATGTTCAATACGGAATAATCC 1260
Db    1201  CGCTTCCATTTTGTATGCTGGGTTTCGTACATGATTCACATGTTCAATACGGAATAATCC 1260
QY    1261  TGATTCTCAAGCTGCCCAACAGGAGCTCGAGCAACAGCTAGACGAGGAAAGCGCTGG 1320
Db    1261  TGATTCTCAAGCTGCCCAACAGGAGCTCGAGCAACAGCTAGACGAGGAAAGCGCTGG 1320
QY    1321  AGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 1380
Db    1321  AGATGACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGG 1380
QY    1381  TAAAGCTGGGCAACAAACAGGGCATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGTGT 1440
Db    1381  TAAAGCTGGGCAACAAACAGGGCATACCTCAATGCTTTAGGACAGATCGCTTCTGCTGTGT 1440
QY    1441  TGTGAGCGCAGAGT 1455
Db    1441  TGTGAGCGCAGAGT 1455
```

RESULT 10

```
US-09-841-132-57
; Sequence 57, Application US/09841132
; Patent No. US20020061948A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 57
; LENGTH: 1537
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-09-841-132-57
```

```
Query Match      8.7%; Score 187.2; DB 9; Length 1537;
Best Local Similarity 53.0%; Pred. No. 1.4e-38;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;

QY    1167  TTTCCATGCTGTAGATGATGCTGAAAATGAGACGCTTCCATTTTGATGCTGGGTTTC 1226
Db    186  TTTCTTTGTTGCTTGATGATGATAGCAATGAATGAGCGAGTTGCAATGCAAGGTTTC 245
QY    1227  GTCAGATGATTCACATGTTCAATACGGAATAATCCCTGATTTCTCAAGCTCCCAACAGGAGC 1286
Db    246  GATCTATGATCGAACAAATTAATGTAAACAATCCCTGCAACAGCTAAAGAGCTACAGCTA 305
QY    1287  TCGAGCAACAGCTAGACGAGCAAGCGCTGGAGATGACAGTGTCTGCTGAGCGGCTGG 1346
Db    306  TGGAGGCTCAGCTGA---CTGCGATGTGATCAACTGTTGGTGGCGATGCGAGCTCC 362
QY    1347  CAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACAAC--- 1397
Db    363  CAGCGCAAAATCAGCAATCAAGATGCTCTTGGCGAAGCTTTGAAACAACATCAGCAG 422
QY    1398  AGGGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTC 1457
Db    423  ATGTTTGTAGCTACAGCTATGGACAGTGGCTTTTGGAGCTGCCNAGGTTGGAGAGGCT 482
QY    1458  CTCCCGCTGCAGCAAGTTCTATAGGCTCATCTGTAAAAACAGCTTTTACAAGACCTCAAAAT 1517
```

Db 483 CGCAGGAACAGCTGGCACTGTCAGATGAATGTAAACAGCTTTACAGACAGCGTTT 542
Qy 1518 CTACAGGTTCT-----GATTATAAACACAGATATCAGCAGGTTATGATGCTTACAAT 1571
Db 543 CTTGACATCTTCCAGCTCTTATGACGAGCACTTTCCGATGGATTTCTGTTACAAA 602
Qy 1572 CCATCAATGATCGCTATGTTAGGCGACGAAATGATCGCACTCGTGATGTATAAACAAATG 1631
Db 603 CACTGAATCTTTATATTCGGAAGCAGAGCGCGTGCACTATAGTCAAACTG 662
Qy 1632 TAAATACCCCGCTCTCAACAGTCCGTTCTTAGACAGCAAGCAAGCTCCAGACACAG 1691
Db 663 CAAATCCCGCGCTTCCAGAGCGTTTCTCGTTCTGGCATAGAAAGCTCAAGGACGCAATG 722
Qy 1692 AAAAAACAGATCAAGCCCTCGTAGGTTGATTTCTGCAATAGCAGAACTCTTGAGATG 1751
Db 723 CAGATGCTAGCAGAGCAGAGAACTATTGTTCAGATAGCCAAAGCTTAGGTGATG 782
Qy 1752 TCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCAATCATCCAGTCAATCTCAAG 1811
Db 783 TATATAGCGCTTACAGGTTCTGGATTTCTGATGCTACGATGTGAGCAATCCGCAAG 842
Qy 1812 CGAATATAGAGATCAGACAAAGCTTACATCGGCACTGACAAAGCTCCACAGTTTG 1871
Db 843 CAAATCAAGAAGAGATTTATGCAAGCTCAGCGCATCTATTAGCAAGCTCCACAAATTTG 902
Qy 1872 GCTATCTTATGTGCAACTTTCTAATGACTCTACACAGAGTTTATAGCTAATATAGAA 1931
Db 903 GGTATCTCTGTTGAGATTTCTGGAATGAGCTTGCAGAAAGTTTCTGCAATTTGAA 962
Qy 1932 GTTGTGTTCTGAAGATCTAGACAGCAGCTGCAATATAAAGCACTTTCTTTGAAACGA 1991
Db 963 CAGAGTTTCTGATCGGGAAGCTAGTCTCGCAATCTCAAGAGATCGTTTAGAAAC 1022
Qy 1992 ACTCTGTTGTTATCAGAGGCTGCTGCTCAATATCGGCTCTATATTCGTTATCTCC 2051
Db 1023 AGCCCGCTTTCATTAACAGGTTGGTAAACATTCCTATCTCTGTTATCTTT 1082
Qy 2052 AATAAC 2057
Db 1083 CTTAAC 1088

RESULT 11

US-09-841-132-412
; Sequence 412, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 21021.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 412
; LENGTH: 1941
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-412

Query Match 8.6%; Score 185.4; DB 9; Length 1941;
Best Local Similarity 48.5%; Pred. No. 5e-36;
Matches 717; Conservative 0; Mismatches 721; Indels 39; Gaps 6;
Qy 606 AAGTTGGCGGCAAAATACAGAAATAGCGAAATATCTTCGGATAACCAAGCGATCTTG 665
Db 467 AGGTTATTGATAGCTCAATCAGTTAGTTAACTTGAAAAACAGATCAGACTTAAAGG 526

Qy 666 ACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTTACAGGCTGCTCTTCTCCAATCTGTAG 725
Db 527 AAACCTTTAAACACACAGACTCTGCAGATCAGATTCAGAGGATTAATAGTCAGTTAGAGA 586
Qy 726 CAAACAAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGGAAGATAACCCAGTAGTCCCGAG 785
Db 587 TCAACAAAAATTTCTGCAGATCAAAATATCAAAAGATCTGGAAGGACAAAACATAAGTTATG 646
Qy 786 GGAACCCCTCGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACACAGA 845
Db 647 AAGCTGTTCTCACTAACCGCAGAGAGGTTATCAAAAGCTTCTTCTGAAGCGGAATTAAGT 706
Qy 846 TAGAGAAAAGATGGAATTCGATTTAGGGATGATATTTTGCAGGACAGAACGCTAGTGGAG 905
Db 707 TAGGACAAAGCTTTCAGTCTATTTGTGATGCTGGGGATCAAAAGCCAGGCTCAGTTCTTC 766
Qy 906 CTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATAGATTCAGCTAAGCAGAA 965
Db 767 AAGCACAGCAAAATTAATAGCCAGATAATATCGCAGCCACGAAGAAATTAATTTGATGCTG 826
Qy 966 TCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCCAATTC 1025
Db 827 CTGAAAACGAGGTAACGAGTTAAACAGAGCATACAGGCTAACGGACTCGCCTTTAG 886
Qy 1026 TTCAAGAACGCGGAACAAATGGTAATACAGGCTGAGAAAGATCTTTAAAAATATCAAACTG 1085
Db 887 TGAATAAAGCTGAGGAGCAGATTAGTCAAGCACAAAAGATATTCAGAGATCAAACT- 945
Qy 1086 CAGATGTTCTGATGTTCCAAATCCAGGAATCAGAGTGGAG-----GCTCAAGCAACA 1140
Db 946 --AGTGGTTCCGATATTCCTATCGTTGTCGAGTGGGTGAGTCTTCCGAGGAAGTG 1003
Qy 1141 AGGAAGTATGATTTGTTAGTAT-----TCGTGTTTCCATGCTGTTTGTAGATGATG 1187
Db 1004 CGTAGAGAGCTTGAAATCTCTAACTTCAAGATCAGNAGAAATTCCTTGTGTTGATG 1063
Qy 1188 CTGAAAATGAGACCGCTTCCAAATTTGATGTTGCGGTTTCGTGAGATGATCAGATGTTCA 1247
Db 1064 TAGACAATGAATGCGCAGGATTCGAATGCAAGGTTTTCGATCTATGATCGAACAAATTA 1123
Qy 1248 ATACGGAATTCCTGATTTCTCAAGCTGCCAACAGGAGCTCGCAGCAAGCTAGACAG 1307
Db 1124 ATGTAAACAATCCTGCAACAGCTAAAGAGCTACAGCTATGGAGCTCAGCTGA---CTG 1180
Qy 1308 CGAAGCCGCTGGAGATGACAGTGTCTGTCAGGCTCGCAGATGCTCAGAAAGCTTTAG 1367
Db 1181 CGATGTGAGATCAACTGTTGTTGCGGATGCGAGCTCCGAGCGAAATACAGCAATCA 1240
Qy 1368 AAGCGCTCTAGGTAAAGCTGGGCAACAAAC-----AGGGCATATCTCAATGTTTATG 1418
Db 1241 AAGATGCTCTTGGCAAGCTTTGAAAACCAACCATCAACAGATGGTTTAGTACAGCTATGG 1300
Qy 1419 GACAGATCGCTCTGCTGCTGTTGTGAGCGCAGAGGTTCTTCCGCTCAGCAAGTTCTTA 1478
Db 1301 GACAAAGTGGCTTTTGACGTGCGAAGTTGGAGGAGGCTCCGAGGAACAGCTGCACTG 1360
Qy 1479 TAGGTCATCTGTAAAAACAGCTTTTACAAGACTCAAAAATCTACAGGTTCTGTA-----TT 1532
Db 1361 TCCAGATGAATGTAAAAACAGCTTTTACAAGACAGCGTTTCTTCGACTTCTTCCAGCTCT 1420
Qy 1533 ATAAACACAGATATCAGCAGGTTATGATGCTTACAAATCCATCAATGATGCTTATGTA 1592
Db 1421 ATGACAGCAGCACTTTCCGATGATTTCTGCTTACAAAACACTGAATCTTTTATATCCG 1480
Qy 1593 GGGCACGAATGATGCGACTCGTGTATGATATAACATGTAAGTACCCCGCTCTCACAC 1652
Db 1481 AAAGCAAGAGCGGCTGTCAGTCACTATTAGTCAAACTGCAAAATCCCGGCTTTCCAGAA 1540
Qy 1653 GATCCGTTCTTAGACAGCAAGAAAGCTCGAGGACCAAGAAAAACAGATCAAGCCCTCG 1712
Db 1541 CGGTTTCTCGTTCTGGCATAGAAAGTCAAGAGCAGCTGCGAGTGTAGTCCAAAGAGCAG 1600
Qy 1713 CTAGGGTGATTTCTGGCAATAGCAGAACTCTTGGAGATGTCTATAGTCAAGTTTGGCAC 1772

Db 1601 CAGAACTATTGTGACAGATAGCCAAAGCTTAGGTGATGATATATAGCGCTTACAGGTTTC 1660
 QY 1773 TACAATCTGTAAATCAGATGATCCAGTCCGAAATCTCAAGCGCAATTAATGAGGATGACAG 1832
 Db 1661 TGAATCTTTGATGCTACGATGTGAGCAATCCGCAAGTAAATCAAGAGAGATTATGC 1720
 QY 1833 AAAAGCTTACATCGGAGTGACAGAGCTCCACAGCTTTGGCTATCTTATGTCGAATTT 1892
 Db 1721 AGAAGCTCACGGATCTCTATTAGCAAGCTCCCAATTTGGGTATCCTGCTGTTCAGAATT 1780
 QY 1893 CTAATGACTCTACACAGAGTTTCATAGCTTAAATTAGAAAAGTTTGTGCTGAAGATCTA 1952
 Db 1781 CTGCGGATAGCTTGAGAAATTTGCTGCGCAATTGGAAGAGAGTTTGTGATGGGAAC 1840
 QY 1953 GGACAGAGCTGAATATAAAGCACTTTCTTTGAAACGAACTTCCTTGTATTATTCAGAGG 2012
 Db 1841 GTAGTCTCGCAGATCTTCGAGAGAATGCGTTTGTAGAAAACAGCCCGCTTTCAATTCACAGG 1900
 QY 2013 TGCTGCTCAATATCGCTCTCTATATCTTCTGTTATCT 2049
 Db 1901 TGTGGAACATGCTTCTCTATCTCTGTTATCT 1937

RESULT 12

US-09-841-260-13

; Sequence 13, Application US/09841260

; Publication No. US20030175700A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Probst, Peter

; APPLICANT: Stromberg, Erika Jean

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS

; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.515

; CURRENT APPLICATION NUMBER: US/09/841,260

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 140

; SEQ ID NO 13

; LENGTH: 1171

; TYPE: DNA

; ORGANISM: Chlamydia trachomatis

US-09-841-260-13

Query Match 8.5%; Score 184; DB 10; Length 1171;
 Best Local Similarity 52.8%; Pred. No. 8.2e-38;
 Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;
 QY 1167 TTCCATGCTGTAGATGATGTAATAATGAGACCGCTTCCATTTTGAATGATGCTGGGTTTC 1226
 Db 207 TTCTCTTGTGCTGATGATGATAGACATGAAATGGCAGCGATTGCCTGCAAGGTTTTC 266
 QY 1227 GTCAGATGATTCACATGTTCAATACGGAATCCCTGATTTCTCAAGTCCCAACAGGAGC 1286
 Db 267 GATCTATGATGAACAATTAATGTAAACAATCCCTGCAACAGCTAAAGAGCTACAAGCTA 326
 QY 1287 TCGCAGCAAGCTAGACAGCGCAAGCCGCTGGAGATCACAGTGTGCTGACGCGCTGG 1346
 Db 327 TGGAGGCTCAGCTGA---CTGCAATGTCAGATCACTAGTTGTTGGGATGGCGAGCTCC 383
 QY 1347 CAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGGTAAAGCTGGGCAACAAAC----- 1397
 Db 384 CAGCCGAAATACAAAGCAATCAAGATGCTCTTGGCGAAGCTTTGAAACAACCATCAGCAG 443
 QY 1398 AGGGCACTCAATGCTTTTAGACAGATCGCTCTGCTGCTGTTGTGAGCGAGGATTC 1457
 Db 444 ATGGTTTGGCTACAGCTATGGGACAGTGGCTTTTTCAGCTGCCAAGGTTGGAGGAGCT 503
 QY 1458 CTCGCCCTGCAGCAAGCTTCTATAGGCTCATCTGTAAACAGCTTTTCAAGAGCTCAAAAT 1517
 Db 504 CCGCAGAACAGCTGCACTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 563
 QY 1518 CTACAGGTTCT-----GATTATAAACAACAGATATCAGACAGGTTATGATGATGATGAT 1571

Db 564 CTTGCACTTCTTCAGCTCTTATGACAGACACTTTCGATGGATATCTGCTTACAAA 623
 QY 1572 CCATCAATGATGCTATGTTAGGSCACGAAATGATGAGCTCGGTGATGATGATGATGATGATG 1631
 Db 624 CACTGAACTCTTTATATTCGAAAGCAGAGCGGCTGCACTAGCTATTAGTCAAACTG 683
 QY 1632 TAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGAGCTCGAGGACGAG 1691
 Db 684 CAAATCCCGCTTTCAGAAAGCGCTTCTCGTTCTGGCATAGAAAGTCAAGGACGAGT 743
 QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGTTGATTTCTGGCAATGACAGAACTCTTTGGAGATG 1751
 Db 744 CAGATGCTAGCCAAAGAGCAGCAGAAACTATTGTACAGAGATAGCCAAAGCTTAGGTGATG 803
 QY 1752 TCTATAGTCAAGTTTCGCACTCAANTCTGTAAAGCAGATCATCCAGTCCGATCTCAAG 1811
 Db 804 TATATAGCGGTTACAGGTTCTGGAATCTTTGATGTTCTACGATTTGTGAGCAATCCGCAAG 863
 QY 1812 CGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTCAAAAGCTCACAAGCTCCACAGTTTG 1871
 Db 864 CAAATCAAGAGAGATTTATGCAAGAGCTCAGCGCATCTATTAGCAAGCTCCCAATTTG 923
 QY 1872 GCTATCCTTATGTCAACTTTCTAATGACTCTACACAGAGTTCAATAGCTAAATAGAAA 1931
 Db 924 GGTATCCTGCTGTTTCAAGATTTCTGGGATAGCTTGCAAGCTTTGCTGCGCAATTTGAAA 983
 QY 1932 GTTTGTTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAGCAGCTTTCCTTTGAAACGA 1991
 Db 984 GAGAGTTTGTGATGGGGAAGCTAGTCTCGCAGAACTCTCAAGAGAACTCGTTTGAAGAAC 1043
 QY 1992 ACTCCTTGTATTATTCAGCAGGCTGCTGCTCAATATCGGCTCTCTATATTCTGTTATCTCC 2051
 Db 1044 AGCCGCTTTCATTCACAGAGTGTGGTAACATGCTTCTCTATCTCTGTTATCTTT 1103
 QY 2052 AATAAC 2057
 Db 1104 CTTAAC 1109

RESULT 13
 US-10-007-693-13
 ; Sequence 13, Application US/10007693
 ; Publication No. US20020146776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515C2
 ; CURRENT APPLICATION NUMBER: US/10/007,693
 ; CURRENT FILING DATE: 2001-12-05
 ; NUMBER OF SEQ ID NOS: 157
 ; SEQ ID NO 13
 ; LENGTH: 1171
 ; TYPE: DNA
 ; ORGANISM: Chlamydia trachomatis
 US-10-007-693-13

Query Match 8.5%; Score 184; DB 13; Length 1171;
 Best Local Similarity 52.8%; Pred. No. 8.2e-38;
 Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;
 QY 1167 TTTCATGCTGTAGATGATGCTGAAAATGAGACCGCTTCCATTTTGTATGTTCTGGGTTTC 1226
 Db 207 TTCTCTTGTGCTGATGATGATAGCAATGAAATGGCAGCGATTGCCATTCGCAAGGTTTC 266
 QY 1227 GTCAGATGATTCACATGTTCAATACGGAATAATCCCTGATTTCTCAAGCTCCCAACAGGAGC 1286
 Db 267 GATCTATGATCGAACTTTAATGTAAACAATCTCTGCAACAGCTAAAGAGCTACAAGCTA 326
 QY 1287 TCGCAGCAAGCTAGACAGCGGAAAGCGCTGGAGATGACAGTGTCTGCTGAGCGCTGG 1346

Db 327 TGGAGGCTCAGCTGA---CTGGATGTCAGATCAACTGGTTGGTGGGATGCGAGCTCC 1383
QY 1347 CAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGTAAAGCTGGGCAACAAC----- 1397
Db 384 CAGCGAAATCAAGCAATCAAGATGCTCTTGGCGAAGCTTTGGAACAACATCAGCAG 443
QY 1398 AGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTTC 1457
Db 444 ATGGTTGGCTACAGCTATGGACAGTGGCTTTTGACGCTCCAGGTTGGAGAGGCT 503
QY 1458 CTCGCGCTCAGCAGAGTTCTATAGGTCATCTGTAAACAGCTTTACAGAGCTCAAAAT 1517
Db 504 CCGCAGGAACAGCTGCGACTGTCCAGATGAATGTAAACAGCTTTACAAGACAGCGTTT 563
QY 1518 CTACAGTTCT-----GATTATAAACACAGATATCAGCAGGTTATGATGTTACAAT 1571
Db 564 CTTCGACTCTTCCAGCTCTTAAGCAGCAGCTTTCCGATGGAATTCGTTTCAAAA 623
QY 1572 CCATCAATGATGCTATGTAGGCGACGAAATGATCGCACTCGTGTATGTATAAACAAATG 1631
Db 624 CACTGAATCTTTATATTCGAAAGCAGAGCGGCTGCAGTCAGCTATTAGTCAAATG 683
QY 1632 TAAGTACCCGCTCTCACACGATCGGTTCTAGAGCAGCAACAGAGCTCGAGGACGAG 1691
Db 684 CAAATCCCGCGCTTTCCAGAACGTTTCTCGTTCTGGCATAGAAAAGTCAAGGACGAGTG 743
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAAGCTCTTGAGATG 1751
Db 744 CAGATGCTAGCCAAAGACGAGCAGAAACTATTTGTGAGATGACCAAGCTTAGTGATG 803
QY 1752 TCTATAGTCAAGTTTGGCAGCTACAACTGTAATGAGATCAATCCAGTCGAATCCTCAAG 1811
Db 804 TATATAGCGCTTACAGGTTCTGGATTTCTTGATGCTACGATTTGAGCAATCCGCAAG 863
QY 1812 CGAATAAGTACAGATCAGAAAGCTTACATGGCAGTGAACAAGCTCCACAGTTTG 1871
Db 864 CAAATCAAGAGAGATATGCAAGAGCTCACGGCATCTATTAGCAAGCTCCACAATTG 923
QY 1872 GCTATCCTTATGCAACTTCTTAATGACTCPACACAGAGTTTATAGTAAATTAGAAA 1931
Db 924 GGTATCCTGCTTTAGAAATCTGGGATAGCTTGCAGAGTTTCTGGCAATTTGAAA 983
QY 1932 GTTGTGCTGAAGATCTAGGACAGCAGCTGAAATAAAGCACTTTCTTTGAAACGA 1991
Db 984 GAGAGTTTGTATGGGAAACGTAGTCTCGCAGAACTTCAAGAGAAATGCTTTAGAAA 1043
QY 1992 ACTCCTGTTTATCAGAGGCTGCTGCTCAATATCGGCTCTCTATATTTCTGTTATCCTC 2051
Db 1044 AGCCGCTTTCAATCAACAGAGTTTGGTAAACATTTCTCTATTTCTCTGTTATCTTT 1103
QY 2052 AATAAC 2057
Db 1104 CTTAAC 1109

RESULT 14
US-09-841-260-38
; Sequence 38, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 38
; LENGTH: 1834
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis

US-09-841-260-38

Query Match 8.5%; Score 184; DB 10; Length 1834;
Best Local Similarity 52.8%; Pred. No. 1.1e-37;
Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;

QY 1167 TTTCATGCTGTAGATGATGTAATGAGACGCTTCCATTTTGTATGTTGTTGTTTC 1226
Db 816 TTTCCTTGTGCTTGTATGATGTAGACAATGAATGGCAGCGATTCATCTGAGGTTTC 875
QY 1227 GTCAGATGATTCAATGTTCAATACGGAAATTCCTGATTCTCAAGCTCCCAACAGGAGC 1286
Db 876 GATCTATGATCGAACAAATTTATGTAAACAATCTCTGCAACAGCTPAAAGAGCTACAGCTA 935
QY 1287 TCGCAGCAAGAGCTTAGACAGCGGAAAGCGGTGGAGATGACAGTCTGCTGAGCGCTGG 1346
Db 936 TGGAGGCTCAGCTGA---CTGCGATGTGAGATCAACTCGTTGGTGGGATGCGAGCTCC 992
QY 1347 CAGATGCTCAGAAAGCTTTAGAAGCGGCTCTAGGTAAAGCTGGGCAACAAC----- 1397
Db 993 CAGCGAAATACAGCAATCAAGATGCTCTTGGCAAGCTTTGAAACAACONTCAGCAG 1052
QY 1398 AGGCATACTCAATGCTTTAGGACAGATGCGCTTCTGCTGCTGTTGTGAGCGCAGGATTC 1457
Db 1053 ATGGTTTGGCTACAGCTATGGGACAAAGTGCGCTTTTGCAAGCTGCCAAGGTTGAGGAGCT 1112
QY 1458 CTCGCGCTGACGCAAGTTCTATAGGTCATCTGTAAACAGCTTTTACAAGACCTCAAAAT 1517
Db 1113 CCGCAGGAAACAGCTGGCACTGTCAGATGAATGTAAACAGCTTTTACAAGACAGCGTTT 1172
QY 1518 CTACAGGTTCT-----GATTATAAACACAGATATCAGCAGGTTATGATGTTACAAT 1571
Db 1173 CTTGCACTTCTTCCAGCTCTTATGCAAGCAGCACTTCCGATGGATATCTGCTTACAAA 1232
QY 1572 CCAATCAATGATGCGCTATGGTAGGCGAAGAAATGATGGAGCTCGTGTATGATAACAATG 1631
Db 1233 CACTGAATCTTTATATTTCCGAAAGCAGAGCGGCTGCAGTCAGCTATTAGTCAAACTG 1292
QY 1632 TAAGTACCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGCTCAGGACGAG 1691
Db 1293 CAAATCCGCGCTTTCCAGAAAGCGTTTCTCGTTCTGGCATAGAAAGTCAAGGACGAGTG 1352
QY 1692 AAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGCAATAGCAGAACTCTTGGAGATG 1751
Db 1353 CAGATGCTAGCCAAAGACAGCAAGAACTATTGTGAGATAGCCAAACGTTAGGTGATG 1412
QY 1752 TCTATAGTCAAGTTTGGCACTACAATCTGTAAATGCGAGATCATCCAGTCGAAATCTCAAG 1811
Db 1413 TATATAGCGGCTTACAGGTTCTGGATTTCTTTGATGTTCTACGATTTGTGAGCAATCCGCAAG 1472
QY 1812 CGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCTCCACAGTTTG 1871
Db 1473 CAAATCAAGAGAGATATGACAGAGCTCAGGCACTCTATTAGCAAGCTCCACAAATTG 1532
QY 1872 GCTATCCTTATGTCAACTTTCTTAATGACTCTACACAGAAAGTTTCATAGCTAAATTAGAAA 1931
Db 1533 GGTATCCTGCTGTTTCAAGATTTCTCGGATAGCTTTGCAAGAGTTTGTCTGGCAATTTGAAA 1592
QY 1932 GTTGTGTTGTAAGGATCTAGGACAGAGCTGAAATAAAGCACTTTCTTTGAAACGA 1991
Db 1593 GAGAGTTTGTGATGGGAAACGTAGTCTCGCAGAACTCTCAAGAGATGCGTTTAGAAAAC 1652
QY 1992 ACTCCTTGTATTACAGAGCTGCTGCTCAATATCGGCTCTCTATATTTCTGTTATCTCC 2051
Db 1653 AGCCGCTTTCAATCAACAGAGTTTGGTAAACATTTCTCTATTTCTCTGTTATCTTT 1712
QY 2052 AATAAC 2057
Db 1713 CTTAAC 1718

RESULT 15
US-10-007-693-38

```

; Sequence 38, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Aay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE TREATMENT OF CHLAMYDIA
; TITLE OF INVENTION: AND DIAGNOSIS
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 38
; LENGTH: 1834
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-007-693-38

```

Query Match	8.5%	Score 184;	DB 13;	Length 1834;
Best Local Similarity	52.8%	Pred. No. 1.le-37;		
Matches 478;	Conservative	0;	Mismatches 410;	Indels 18; Gaps 3;
QY	1167	TTTCCATGCTGTTAGATGATGCTGAAATCAGACCGCTTCATTTGATGCTCGGGTTTC	1226	
Db	816	TTTCTTTGTTGTTGATGATGTAGCAATGGAATGGCAGGATTCAGCTCAAGGTTTC	875	
QY	1227	GTCCAGATGATTCACATGTTTCAATACGGAAATCCTGATTTCAAGTCGCCCAACAGGAGC	1286	
Db	876	GATCTATGATCGAACAAATTTAATGTAAACAATCCTGCAACAGCTAAAGAGCTACAAGCTA	935	
QY	1287	TGCGACACAAGCTAGACGACGCAAGCCGCTGGAGATGACAGTGTCTGTGACGCGCTGG	1346	
Db	936	TGGAGGCTCAGCTGA---CTGCCATGTCAATCAACTGGTTGGTGGGATGGCAGAGCTCC	992	
QY	1347	CAGATGCTCAGAAAGCTTTAGAACGGCTCTAGCTAAAGCTGGGCAACAAC-----	1397	
Db	993	CAGCCGAATACAAAGCAATCAAGATGCTCTTGGCAGAGCTTTGAAACAACCATCAGCAG	1052	
QY	1398	AGGGCATCTCAATGCTTTAGGACAGATCGCTCTCTGCTGTGTTGTGAGCGCAGGAGTTC	1457	
Db	1053	ATGGTTTGGCTACAGCTATGGGACAAAGTGGCTTTTGACGTGCGCAAGGTTGGAGGAGGCT	1112	
QY	1458	CTCCGCTCAGCAAGTCTCTATAGGTCATCTCTGTAACAGCTTTACAGACCTCAAAAT	1517	
Db	1113	CGCAGGAACAGCTGGCACTGTCCAGATGATGTAAACAGCTTTACAGACAGCGTTTT	1172	
QY	1518	CTACAGGTTTCT-----GATTATAAAACACAGATATCAGCAGGTTATGATGCTTACAAAT	1571	
Db	1173	CTTCGACTTCTCCAGCTCTTATGCAGCAGCACTTCCGATGGATATCTGCTTACAAA	1232	
QY	1572	CCATCAATGATGCTATGTTAGGCGCAAAATGATCGGACTCTGTGATGTGATATAACAATG	1631	
Db	1233	CAGTGAACCTCTTTATATTCGAAAGCAGAGCGGCTGCACTCAGCTATTAGTCAAACGT	1292	
QY	1632	TAAGTACCCCGCTCTCACCATCGCTTCTTAGACACGAAACAGAGCTCGAGGACCAG	1691	
Db	1293	CAAAATCCCGCGCTTCCAGAGCGTTTCTCGTCTGGCATAGAACTCAAGGACGAGTG	1352	
QY	1692	AAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCTTTGGAGATG	1751	
Db	1353	CAGATGCTAGCCAAAGAGCAGCAAAACTATTGTCAAGATAGCCAAAACGTTAGGTGATG	1412	
QY	1752	TCTATAGTCAAGTTTTCGGCACATCAATCTGTANTGCAGATCATCCAGTCAATCTCAAG	1811	
Db	1413	TATATAGCGCTTACAGGTTCTGGATTCCTTTGATGTTCTAGATGTGAGCAATCCCGAAG	1472	
QY	1812	CGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCAGTGAACAAAGCCCTCCACAGTTG	1871	
Db	1473	CAAAATCAAGAGAGATTATGCAGAAAGCTCAGCGCACTATTAGCAAGCTCCACAATTC	1532	
QY	1872	GCTATCGTTATGCGAATCTTAAATGACTCTACAGAAAGTCTATAGCTAAATTAGAAA	1931	
Db	1533	GGTATCCTGCTGTTTCAGAAATCTTCGGGATAGCTTGCAGAAAGTTGCTGGCAATTCGAAA	1592	

Qy	1932	GTTTGTTT	GCTGAAGGATCT	TAGGACAGCAGCT	GAATAAAAGCAC	TTTCCTTTGAACGA	1991
Db	1593	GAGNGTTT	GTGATGGGAACGT	TAGTCTCGCAGAA	TCTCAAGAGAA	TGCGTTTAGAAAC	1552
Qy	1992	ACTCTCTG	TTTATTACAGCAGG	TCTGGTCAATA	TTCGGCTCTCTATAT	TTCTGGTTATCTCC	2051
Db	1653	AGCCCGGT	TTCAITTCACAGG	TCTGGTAAACAT	TGCTCTCTATTTCT	CTGGTTATCTTT	1712
Qy	2052	AATAAC	2057				
Db	1713	CTTTAAC	1718				

Search completed: March 23, 2004, 19:25:08
Job time : 757.183 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:23:38 ; Search time 160.201 Seconds
(without alignments)
7067.770 Million cell updates/sec

Title: US-10-608-559-1

Perfect score: 2156

Sequence: 1 ataaatctttaaaacagg.....gatctcaataataaaga 2156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1155024 seqs, 262583689 residues

Total number of hits satisfying chosen parameters: 2310048

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_New.*

1: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/prodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/prodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/prodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/prodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	2.3	1860	6 US-10-414-532-71	Sequence 71, Appl
2	49	2.3	2000	6 US-10-414-532-20	Sequence 20, Appl
3	45.8	2.1	775	6 US-10-414-532-29	Sequence 29, Appl
4	45.8	2.1	1473	6 US-10-414-532-33	Sequence 33, Appl
5	41.2	1.9	416	6 US-10-767-701-16328	Sequence 16328, A
6	41	1.9	234309	6 US-10-796-280-12378	Sequence 12378, A
7	40	1.9	1790242	6 US-10-767-471-10805	Sequence 10805, A
8	39.6	1.8	447	6 US-10-767-701-17775	Sequence 17775, A
9	39.6	1.8	466	6 US-10-767-795-178	Sequence 178, App
10	39.2	1.8	165223	6 US-10-796-280-12544	Sequence 12544, A
11	38.8	1.8	335	6 US-10-793-479-31955	Sequence 31955, A
12	38.8	1.8	318752	6 US-10-767-471-10553	Sequence 10553, A
13	38.4	1.8	52746	6 US-10-765-790-39	Sequence 39, Appl
14	38.4	1.8	119006	7 US-60-550-051-3011	Sequence 3011, Ap
15	38.2	1.8	2685	6 US-10-784-480-357	Sequence 357, App
16	38	1.8	770	1 PCT-US04-05757-19	Sequence 19, Appl
17	38	1.8	770	6 US-10-776-889-19	Sequence 19, Appl
18	38	1.8	967	1 PCT-US04-05757-18	Sequence 18, Appl
19	38	1.8	967	6 US-10-776-889-18	Sequence 18, Appl
20	38	1.8	1160	1 PCT-US04-05757-17	Sequence 17, Appl
21	38	1.8	1160	6 US-10-776-889-17	Sequence 17, Appl
22	38	1.8	1367	1 PCT-US04-05757-16	Sequence 16, Appl
23	38	1.8	1367	6 US-10-776-889-16	Sequence 16, Appl
24	38	1.8	1553	1 PCT-US04-05757-15	Sequence 15, Appl
25	38	1.8	1553	6 US-10-776-889-15	Sequence 15, Appl
26	38	1.8	1719	1 PCT-US04-05757-14	Sequence 14, Appl

```

RESULT 1
US-10-414-532-71
; Sequence 71, Application US/10414532
; GENERAL INFORMATION:
; APPLICANT: CURTISS III, ROY
; APPLICANT: KANG, HO YOUNG
; TITLE OF INVENTION: IMPROVED IMMUNOGENIC COMPOSITIONS AND VACCINES COMPRISING
; FILE REFERENCE: 56029-40437
; CURRENT APPLICATION NUMBER: US/10/414,532
; PRIOR FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/372,710
; PRIOR FILING DATE: 2002-04-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 71
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1857)
US-10-414-532-71
Query Match 2.3%; Score 49; DB 6; Length 1860;
Best Local Similarity 49.5%; Pred. No. 0.0073;
Matches 155; Conservative 0; Mismatches 155; Indels 3; Gaps 1;
QY 845 ATAGAGAAAGATGGAATGCCGATTAGGATGCAATATTTTGCAGACAGAGCGCTAGTGGGA 904
DB 346 ACAGACAAAGCGGCAAAAGACGACAGATAGATGATAGATGAAGCTAAGAAACGGAA 405
QY 905 GCTGTAGAAATGCTTAATCTAATAACAGTATAGCAACATAGATTAGCTTAAGCAGCA 964
DB 406 GAAAGGCAAAACATAAATTTAATCTGTTGCGCAATGGTAGTTCCTGAGCCAGCAG 465
QY 965 ATCGTCTACTGTAAAGACACAATAAGCTGAAGCTCAGAAAAAGTTCCCGGA---CTCTCCA 1021
DB 466 TTGCTGAGACTAAGAAAAAATCAGAGAGAGCTAATAAAGCACCAGACACTTACTAAA 525
QY 1022 ATCTTCAAGAGCGGAAACAAATGGTAAATACAGCTGAGAAAGATCTTAAAAATATCAAA 1081
DB 526 AAACCTAGAGAGAGCTAAAGCAAAAATTAGAGAGGCTGAGAAAAAGCTACTGAAGCCAAA 585
QY 1082 CCTGCAGATGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGAGGCTCCAGCAACAA 1141
DB 586 CAAAAGTGGATGCTGAAGAAAGTCGCTCCTCAAGCTAAATCGCTGAATTTGAAATCAA 645
QY 1142 GGAAGTAGTATTG 1154

```

ALIGNMENTS


```
QY 845 ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTCAGGACAGAACGCTAGTGA 904
Db |||||
QY 943 ACAGACAAAGCGCGAAGACCCACAGATAGATGATGATGAAGCTTAAGAACGCGAA 1002
Db |||||
QY 905 GCTGTAGAAAATGCTTAATCTTAATTAACAGTATAACACATAGATTCAGCTAAAGACGA 964
Db |||||
QY 1003 GAAGAGGCAAAATCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1062
QY 965 ATCGCTACTGCTTAAGACACAAATCTGAGCTCAGAAAAGTTCCCGA---CTCTCCA 1021
Db |||||
QY 1063 TTGGCGAGACTTAAGAAAATCTAAGAAAGCTTAACAAAAGACCCAGAACTTACTAAA 1122
QY 1022 ATTCTTCAAGAGCGGAACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAATATCAA 1081
Db |||||
QY 1123 AAACCTGGAAGAGCTTAAGCAAAATTAAGAGGCTGAGAAAAGCTACTGAAGCCAAA 1182
QY 1082 CTGCGAGATGGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAAA 1141
Db |||||
QY 1183 CAAAAGTGGATGCTGGAAGAAAGTCCTCCTCAAGCTAAATCGCTGAATTGAAAAATCAA 1242
QY 1142 GGAAGTAGTATTG 1154
Db |||||
QY 1243 GTTCAATGCTCG 1255
Db |||||

RESULT 5
US-10-767-701-16328/c
; Sequence 16328, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 16328
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(416)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3480-022-P1-K1-Fl1
; US-10-767-701-16328
```

```
Query Match 1.9%; Score 41.2; DB 6; Length 416;
Best Local Similarity 62.7%; Pred. No. 0.4;
Matches 64; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1 ATAAAACTTTAAACAGGCTCGCATTAATATAGTGAAGCTTTTTTTTATTTTT 60
Db |||||
QY 61 ATATAAACTTAAGATTTTATATATTTTGTAGTTTAT 102
Db |||||
QY 161 ATTTTAAATAATTTTATTTTATTTATATATATATAT 120
Db |||||
```

```
RESULT 6
US-10-796-280-12378/c
; Sequence 12378, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001510
; CURRENT APPLICATION NUMBER: US/10/796,280
```

```
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12378
; LENGTH: 234309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(234309)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-796-280-12378

Query Match 1.9%; Score 41; DB 6; Length 234309;
Best Local Similarity 52.0%; Pred. No. 7.5;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 913 AAATGCTAAATCTAATTAACAGTATAAGCAACATAGATTCAGCTAAAGCAGCAATCGCTAC 972
Db |||||
QY 973 TGCTAAGACACAAATAGCTGAGCTCAGAAAAGTTCCCGACTCTCCAATTTCTAAGA 1032
Db |||||
QY 155018 AGTGAAGAGACAAACCCACAGAAATGGGAGAAAATATTTGCAACCATCCATCTGATAAAG 154959
QY 1033 AGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTTAAAAATATCAAACTCGAGA 1089
Db |||||
QY 154958 AACGTAACCACTATACAGGAGCTCAAGAACTCAACAGGAAAAAACTGAAA 154902
Db |||||

RESULT 7
US-10-767-471-10805
; Sequence 10805, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10805
; LENGTH: 1790242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1790242)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-767-471-10805

Query Match 1.9%; Score 40; DB 6; Length 1790242;
Best Local Similarity 51.4%; Pred. No. 32;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 710 CTCTCCAAATCTGTAGCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGAT 769
Db |||||
QY 770 AACCCAGTAGTCCCGAGAAAACGCCCTGCAATTTGCTCAATTTTAGTTGATCAGACAGAT 829
Db |||||
QY 1751671 TAACTGGAATATTTATAGAGAAAAGACCGGAAGAGATGATCAAGTGAAGGAGAGAA 1751730
QY 830 GCTACACGACACAGATAGAGAAAGATGGAATGCCATAGGATGCATATTTTCAGGA 889
Db |||||
QY 1751731 GCAAAAAGAAATGAAGAAAGGAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1751790
QY 890 CAGA 893
Db |||||
QY 1751791 AAAA 1751794
Db |||||
```

RESULT 8
US-10-767-701-17775
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 17775
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(165223)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-767-701-17775

Query Match 1.8%; Score 39.6; DB 6; Length 447;
Best Local Similarity 60.0%; Pred. No. 1.1;
Matches 66; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 6 ATCTTTAAACAGCGCTCGCATTAATTATTAGTGAGAGCTTTTATTTTATTTTATTTTATAAT 65
Db 324 ATTTTAAAAATTTAATGTTTTTTTATATATATATTTTATTTTATTTTATTTTATTTTAT 383
Qy 66 AAAAATAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 115
Db 384 ATTTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 433

RESULT 9
US-10-767-795-178
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 178
; LENGTH: 466
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(165223)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-767-795-178

Query Match 1.8%; Score 39.6; DB 6; Length 466;
Best Local Similarity 63.8%; Pred. No. 1.1;
Matches 60; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 22 TCGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 81
Db 265 TTGTTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 324
Qy 82 TATTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 115
Db 325 TATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTATTT 358

RESULT 10
US-10-796-280-12544
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SEQ ID NO 12544
; LENGTH: 165223
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(165223)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-796-280-12544

Query Match 1.8%; Score 38.8; DB 6; Length 335;
Best Local Similarity 54.1%; Pred. No. 1.5;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 931 CAGTATAAGCAACATAGATTTCAGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGC 990
Db 124 CAGTTTAAACCTTTCTGCAAGCAAAAGAAACAATCAACAAGTGAAGACACACCCAC 183

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12544
; LENGTH: 165223
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(165223)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1.
US-10-796-280-12544

Query Match 1.8%; Score 39.2; DB 6; Length 165223;
Best Local Similarity 47.5%; Pred. No. 1.9;
Matches 116; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 1022 ATCTTCAAGACGGGACAAATGGTAAATACAGCGCTGAGAAAGATCTTAAAAATATCAAA 1081
Db 145793 ATCTTCTTGTATGTGAATGAAAAGATGCACCTTGGAAACTGCTAGTAGGCATTTCAAA 145852
Qy 1082 CCTGCAGATGGTCTGATGTTCCAAATCCAGGAATACAGTTGGAGGCTCCAAAGCAACAA 1141
Db 145853 TGTAAATGCCATTTTAAATATATATAATAGGATTTATATTTATTTTGTACTTATAA 145912
Qy 1142 GGAAGTAGTATTTGATAGTATTCCTGTTTCATGCTGTAGATGATGCTGAAAAATGAGAC 1201
Db 145913 ATAGTATTATTTTGAATTTATTTTCTCATGAAGTTGTTTGTGCAAGTATTGAGAAA 145972
Qy 1202 GCTTCATTTTGTGCTGGTTTCGTCAGATGATTCACATGTTCAATACGGAAATCCT 1261
Db 145973 TTATCCATATTTCTATGTTAATCTTGACATGGGATTTTACTTCTTTTGTAGCAAAAAGT 146032
Qy 1262 GATT 1265
Db 146033 GAAT 146036

RESULT 11
US-10-793-479-31955
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/10/793,479
; CURRENT FILING DATE: 2004-03-03
; PRIOR APPLICATION NUMBER: US/09/513,999
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 31955
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-793-479-31955

Query Match 1.8%; Score 38.8; DB 6; Length 335;
Best Local Similarity 54.1%; Pred. No. 1.5;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 931 CAGTATAAGCAACATAGATTTCAGCTAAGCAGCAATCGCTACTGCTAAGACACAAATAGC 990
Db 124 CAGTTTAAACCTTTCTGCAAGCAAAAGAAACAATCAACAAGTGAAGACACACCCAC 183

```
QY 991 TGAAGCTCAGAAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAATGGTAAT 1050
Db 184 AGAATGTGAAAAATATTCGCAACTAGCCATCTGACAGGGATTAATACCAAGATAT 243
QY 1051 ACAGGCTGAGAAAGATCTTAAATA 1076
Db 244 ACAAGCTTAACAGCTCTATAGAAA 269

RESULT 12
US-10-767-471-10553/c
; Sequence 10553, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10553
; LENGTH: 318752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(318752)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10553

Query Match 1.8%; Score 38.8; DB 6; Length 318752;
Best Local Similarity 62.2%; Pred. No. 32;
Matches 61; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 30 ATTATTAGTCAGAGCTTTTATTTTATTTTATATATAATAAACTAAAGATTTTATTT 89
Db 97038 ATGTGTAATATATCTTCTGTTATTTATTTATTTATTTATTTATTTATTTATTT 96979
QY 90 TTTGAGTTTTATGGTTAATCCTATTGGTCCAGGTCCT 127
Db 96978 TTATGTTATTATCCTTGTTATGCTACTCAAGGCCT 96941

RESULT 13
US-10-765-790-39
; Sequence 39, Application US/10765790
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 52746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-39

Query Match 1.8%; Score 38.4; DB 6; Length 52746;
Best Local Similarity 48.2%; Pred. No. 18;
Matches 108; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
```

```
QY 936 TAAGCAACATGATTTCAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAG 995
Db 26475 TCAGATCCCGGCTATATTAATAGTGAAGACCTCTTACTGAGAAAGACACTTGCAAA 26534
QY 996 CTCAGAAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAATGGTAAATACAGG 1055
Db 26535 GGCCAGAAAAAGGTGGCCATTTCTTCAATGTGAGGACCAACACAAAGACAGAAGATT 26594
QY 1056 CTGAGAAAGATCTTAAATAATATCAACCTGAGAGTGTCTGATGTTTCAATCCAGAA 1115
Db 26595 ATTTAAAAATCAAGAAATATGACACCAAAAGAACTAAATAATGCTCGGTAATAGACC 26654
QY 1116 CTACAGTTGGAGGCTCCCAAGCAACAAGGAAGTAGTATTGGTAGT 1159
Db 26655 CTGAATATGCAAGCTATGAGACCACCTAAGAATTTAAGATAAT 26698

RESULT 14
US-60-550-051-3011
; Sequence 3011, Application US/60550051
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001508
; CURRENT APPLICATION NUMBER: US/60/550,051
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 23014
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3011
; LENGTH: 119006
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-550-051-3011

Query Match 1.8%; Score 38.4; DB 7; Length 119006;
Best Local Similarity 48.2%; Pred. No. 26;
Matches 108; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 936 TAAGCAACATGATTTCAGCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAG 995
Db 31269 TCAGATCCCGGCTATATTAATAGTAGAAGACCTCTTACTGAGAAAGACACTTGCAAA 31328
QY 996 CTCAGAAAAAGTCCCGACTCTCCAAATCTTCAAGAGCGGAAACAATGGTAAATACAGG 1055
Db 31329 GGCCAGAAAAAGGTGGCCATTTCTTCAATGTGAGGACCAACACAAAGACAGAAGATT 31388
QY 1056 CTGAGAAAGATCTTAAATAATATCAACCTGAGATGGTCTGATGTTTCAATCCAGAA 1115
Db 31389 ATTTAAAAATCAGAAAAATATGACACCAAAAGAACTAAATAATGCTCGGTATAGACC 31448
QY 1116 CTACAGTTGGAGGCTCCCAAGCAACAAGGAAGTAGTATTGGTAGT 1159
Db 31449 CTGAATATGCAAGCTATGAGACCACCTAAGAATTTAAGATAAT 31492

RESULT 15
US-10-784-480-357
; Sequence 357, Application US/10784480
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter
; FILE REFERENCE: 06132/043003
; CURRENT APPLICATION NUMBER: US/10/784,480
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US 09/895,913
; PRIOR FILING DATE: 2001-06-29
```

```
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
; LENGTH: 2685
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2601)
; US-10-784-480-357

Query Match      1.8%; Score 38.2; DB 6; Length 2685;
Best Local Similarity 45.7%; Pred No. 5.5; 248; Indels 12; Gaps 2;
Matches 219; Conservative 0; Mismatches 0;

QY 602 GTTAAAGTTGGCGCAATACAGATTAGCGAAATATGCTTCGGATAACCAAGCGATT 661
Db 1327 GATAGTGAAGGGCGCTAATCTAGGAATTAACGACAAGCTGATGAAGTGTGATCTGATT 1386
QY 662 CTGACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCCAACTCT 721
Db 1387 TATACGATCTAAGAAACCGTAAACCTTTGATAACATGGCGGCTAAAGGTTATCCATTG 1446
QY 722 GTAGCAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTC 781
Db 1447 TTACCAATGGATTTCAAAANTGGCGCGATATTGCCACTATTAGCCCACTAATGTTGAT 1506
QY 782 CCAGGGAACCGCCTGCAATTGCTCAATCTTTAGTTGATTCAGACAGATGCTACAGCGACA 841
Db 1507 GCGGACAAAA---TAGCTAGCGATAATCCTATTATTATGCTTCCATAGAGCTGATATTGCC 1563
QY 842 CAGATAGAGAAAGATGGAAATCGGATTTAGGGATGCAATTTTTCAGGACAGAACGCTA-- 899
Db 1564 AAGCAATACGAACACAGAAAAACCAATTAGGATAAGATTTAGAGCTTAATTAGCTAAG 1623
QY 900 -----GTGGAGCTGTAGAAAAATCTAAATCTAATAACAGTATAGCAACATAGATTCA 952
Db 1624 GCTTTAGGTGGCAATAAAAAAGATGACGATAAAGAAAAAGTAAAAAATCCACAGCAGAA 1683
QY 953 GCTAAAGCAGCAATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCC 1012
Db 1684 GCTAAAGCAGAAAAACAATAAGATAGACAAAGATGTCGCAAGAACTGCCAAGATATCAGT 1743
QY 1013 GACTCTCAATTTCTTCAAGAGCGGACAAATGGTAAATACAGGCTGAGAAAGATCTTAA 1071
Db 1744 GAAATCGCTCTTAAGAACAAAAAAGAGAGTGGGGATTTGTAGATGMAAATCGTAA 1802
```

Search completed: March 23, 2004, 18:53:26
Job time : 168.201 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:19 ; Search time 5652.79 Seconds

(without alignments)
11389.568 Million cell updates/sec

Title: US-10-608-559-1

Perfect score: 2156

Sequence: 1 ataaatctttaaacagg.....gatcttcaataataaaga 2156

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1:	em_estba.*
2:	em_esthum.*
3:	em_estin.*
4:	em_estmu.*
5:	em_estov.*
6:	em_estpl.*
7:	em_estro.*
8:	em_htc.*
9:	gb_est1.*
10:	gb_est2.*
11:	gb_htc.*
12:	gb_est3.*
13:	gb_est4.*
14:	gb_est5.*
15:	em_estfun.*
16:	em_estom.*
17:	em_esthum.*
18:	em_estinv.*
19:	em_gss_pln.*
20:	em_gss_vrt.*
21:	em_gss_fun.*
22:	em_gss_mam.*
23:	em_gss_mus.*
24:	em_gss_pro.*
25:	em_gss_rod.*
26:	em_gss_phg.*
27:	em_gss_vrl.*
28:	gb_gss1.*
29:	gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	52.8	2.4	843	28	BZ643413
C 2	52	2.4	1201	9	AL541034
C 3	50	2.3	1201	13	BX446742
C 4	49	2.3	909	13	BX453105

C 5	48.4	2.2	1191	14	CD500670
C 6	48.2	2.2	533	13	BX353273
C 7	48	2.2	872	14	CK097864
C 8	47.8	2.2	510	28	BH714761
C 9	47	2.2	808	29	CG815585
C 10	46.8	2.2	1201	13	BX356412
C 11	46.6	2.2	359	9	AL840303
C 12	46.6	2.2	1201	9	AL531906
C 13	46.6	2.2	1202	28	CC262481
C 14	46.2	2.1	598	14	CD661508
C 15	46	2.1	427	13	BX403499
C 16	46	2.1	1041	9	AL575932
C 17	46	2.1	1201	9	AL558113
C 18	45.8	2.1	642	13	BX707012
C 19	45.8	2.1	1104	13	BX396222
C 20	45.8	2.1	1201	9	AL561244
C 21	45.6	2.1	967	28	BH184682
C 22	45.6	2.1	967	29	CNS07PKF
C 23	45.6	2.1	1201	13	BX381961
C 24	45.4	2.1	528	12	BJ340195
C 25	45.4	2.1	724	12	BJ344100
C 26	45.4	2.1	819	28	BZ391755
C 27	45.4	2.1	1201	9	AL524496
C 28	45.2	2.1	336	14	CB098017
C 29	45.2	2.1	817	14	CD656959
C 30	45.2	2.1	856	13	BUL75987
C 31	45.2	2.1	890	13	BQ424404
C 32	45.2	2.1	1004	13	BX334499
C 33	45.2	2.1	1072	13	BX359378
C 34	45	2.1	1164	28	CC261365
C 35	44.8	2.1	484	9	AA314486
C 36	44.8	2.1	1124	13	BX436282
C 37	44.8	2.1	1151	10	BE964820
C 38	44.8	2.1	1173	9	AL551058
C 39	44.8	2.1	1183	13	BX379850
C 40	44.6	2.1	259	14	CB722003
C 41	44.6	2.1	1537	12	B1195792
C 42	44.4	2.1	856	28	BH132709
C 43	44.4	2.1	885	13	BX425603
C 44	44.4	2.1	976	13	BX388320
C 45	44.4	2.1	1096	13	BX381337

ALIGNMENTS

RESULT 1
BZ643413/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BZ643413
OGAOW57TM ZM.0.7 1.5 KB Zea mays genomic clone ZMMBMA011J17,
genomic survey sequence.
BZ643413
BZ643413.1 GI:28104915
GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 843)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGAOW57TC
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR

CD500670 CDA46-D01
BX353273 BX353273
CK097864 UB64CP002
BH714761 BOMGV71TR
CG815585 SOYFPH83TV
BX356412 BX356412
AL840303 AL840303
AL531906 AL531906
CC262481 CH261-167
CD661508 ETEST6948
BX403499 BX403499
AL575932 AL575932
AL558113 AL558113
BX707012 BX707012
BX396222 BX396222
AL561244 AL561244
BH184682 025_N.14-
AL621633 T3 end of
BX381961 BX381961
BJ340195 BJ340195
BJ344100 BJ344100
BZ391755 EINBH26TR
AL524496 AL524496
CB098017 ku48b04.Y
CD656959 AGENCOURT
BUL75987 AGENCOURT
BQ424404 AGENCOURT
BX334499 BX334499
BX359378 BX359378
CC261365 CH261-72H
AA314486 EST186534
BX436282 BX436282
BE964820 601658475
AL551058 AL551058
BX379850 BX379850
CB722003 jnn603F04
B1195792 602754186
BH132709 ENTNG59TF
BX425603 BX425603
BX388320 BX388320
BX381337 BX381337

```

Class: sheared ends.
Location/Qualifiers
1..843
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBMA011J17"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBOSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      2.4%; Score 52.8; DB 28; Length 843;
Best Local Similarity 47.8%; Pred. No. 1.7;
Matches 153; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 286 TTCTGCAGTGAATGCTCTCATGAGCTGCGAGATAAGCTGGTATGCTTCTAGTAACAG 345
Db 742 TACTGCTACTACTGCTACTGCTGCTACTGCTACTACTACTACTACTACTACTACTGCTAC 683

QY 346 CTCGCTTCTACTAGCAGCTGCGAGAGCTGAGCTCAACGACGAGCGACCGACCTACGCC 405
Db 682 TACTGCTACTACTGCTACTGCTGCTACTACTACTACTACTACTACTACTGCTACTGCTGC 623

QY 406 TCCTCCACGCGTTGATGATTATAGACTCAAGCGCAACAGCTTACGATATCTT 465
Db 622 TACTGCTGCTACTGCTGCTATGCTGCTACTCTCCAGCTACTACTGCTGTTGCTGCTACTGC 563

QY 466 TACCTCAACATCAGTACGCTGACATACAGCTGCTTTGGTGAGCTCCAGGATGCTGCAC 525
Db 562 TGCTGCTACTGCTACTGCTGCTACTGCTGCTGCTACTACTGCTGCTACTGCTACTGCTGC 503

QY 526 TAATAAAGGATACAGCGGCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAAC 585
Db 502 TACTACTACTGCTACTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTACTGCTACTGCTAC 443

QY 586 TAAGATCCGATGAGTTA 605
Db 442 TACTACTACTGCTGCTGCTA 423

RESULT 2
AL541034/c      1201 bp      mRNA      linear      EST 12-MAY-2003
LOCUS          AL541034 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YA21
DEFINITION     3-PRIME, mRNA sequence.
ACCESSION      AL541034
VERSION        AL541034.2 GI:30544815
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL         Full-length cDNA libraries and normalization
COMMENT         Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12871720.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 951.r
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE005A11NP1&cluster=951.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005A11NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YA21"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YA21"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match      2.4%; Score 52; DB 9; Length 1201;
Best Local Similarity 42.0%; Pred. No. 2.4;
Matches 63; Conservative 35; Mismatches 52; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTGAAGCTTTTTTTTTTTTTTTT 60
Db 1160 AAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTATTTTATTTTTTTTTTT 1101

QY 61 ATATAAACTTAAAGATTTTTTATTTATTTTTCAGTTTTTATGTTATCTATGTTCC 120
Db 1100 WAAAAAATAAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCAGTTT 1041

QY 121 AGTCTCTATAGACGAACAGACGACACC 150
Db 1040 VMMMMMMVMMGGGMCAMACRACARRAC 1011

RESULT 3
BX446742/c      1201 bp      mRNA      linear      EST 22-MAY-2003
LOCUS          BX446742 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
DEFINITION     CL0BB011ZG07 5-PRIME, mRNA sequence.
ACCESSION      BX446742
VERSION        BX446742.1 GI:31025761
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL         Full-length cDNA libraries and normalization
COMMENT         Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8247.f
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CL0BB011ZG07RPI&cluster=8247.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CL0BB011ZG07RPI.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BB011ZG07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

```



```

Query Match      2.3%; Score 50; DB 13; Length 1201;
Best Local Similarity 33.9%; Pred. No. 6.2;
Matches 76; Conservative 57; Mismatches 91; Indels 0; Gaps 0;

Qy 7 TCTTTAAACAGAGCTCCATATATAGTAGAGAGCTTTTATTTTATTTTATAATA 66
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1159 TTTTAAATWMMWAWTYKTMMAAYTYRCVGGTTTTTTTTTTTTTTTTTMEHTTMMHW 1100

Qy 67 AAACTRAAAGATTTTATTTATTTTGTGTTTATGTTTATGTTTATCTTATGTCAGGTC 126
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1099 WWWTTWTTTHTTTTTTWTATTTTGTGTTTTCMAWTTTTCMMMAAAAAATTTWKTHM 1040

Qy 127 TATAGACGAACAGACACACCTCCGACAGATCTTTCTGCTCAAGGATTGGAGGCGAG 186
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
1039 MMMAAAAAAMWMAHHHHHHHSSSSGGGRRKKKRGGGGGGGGGGGGGGGGG 980

Qy 187 TGCAGCAATAAGAGTGGGAGAGCTCAAGAATAGCAGGTGCGG 230
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
979 GRRAAVAHAHAHAARGRCMCTTGGAHAHAHAAGGGGGGGGG 936

RESULT 4
BX453105/c
LOCUS BX453105 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone EST 22-MAY-2003
DEFINITION CS0DF014YP19 5-PRIME, mRNA sequence.
ACCESSION BX453105
VERSION BX453105.1 GI:31026455
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 909)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7159.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAH008ZB02QP1&cluster=7159.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAH008ZB02QP1.
FEATURES
source
1. .909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF014YP19"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT.6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match      2.3%; Score 49; DB 13; Length 909;
Best Local Similarity 45.0%; Pred. No. 10;
Matches 94; Conservative 15; Mismatches 100; Indels 0; Gaps 0;

Qy 27 TTAATTATAGAGAGCTTTTATTTTATTTTATATAAATAAAGATTTTATTA 86
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
823 TTTTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 764

```

```

Qy 87 TTTTGTGAGTTTATGTTTAACTTATGTCAGGTCCTATAGACGAACAGCA 146
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
763 TTTTGTGAGTTTATGTTTAACTTATGTCAGGTCCTATAGACGAACGAACGCA 704

Qy 147 CACCTCCCGCAGATCTTTCTGCTCAAGATGAGGCGAGTCGACGAAATAAGAGTCGG 206
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
703 CCCCCNNNNNNNNNTTTTGGAGCAGTCGATGAGTCTTTAATTGGAAGCAA 644

Qy 207 AAGTCAAGAAATAGCAGGTGCGGAAGCT 235
Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
643 WAGKTCATTTAATGAAGGCGAGGCT 615

RESULT 5
CD500670/c
LOCUS CD500670 1191 bp mRNA linear EST 12-JUN-2003
DEFINITION CDA46-D01.yid-s SHGC-CDA Gasterosteus aculeatus cDNA clone
ACCESSION CD500670
VERSION CD500670.1 GI:31427701
KEYWORDS EST.
SOURCE Gasterosteus aculeatus (three spined stickleback)
ORGANISM Gasterosteus aculeatus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
TITLE Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
COMMENT Gasterosteidae; Gasterosteus.
1 (bases 1 to 1191)
Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmutz, J., and Myers, R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
Contact: Kingsley, DM
HMI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B306, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cmgm.stanford.edu
Plate: 45
High quality sequence start: 12
High quality sequence stop: 866.
FEATURES
Location/Qualifiers
source
1. .1191
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas river, CA"
/db_xref="taxon:69293"
/clone="CDA46-D01"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
/clone_lib="SHGC-CDA"
/notes="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI
(5' adaptor); Site 2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dt sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted into the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of pBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual pBK-CMV phagemid clones for
EST sequencing."

ORIGIN
Query Match      2.2%; Score 48.4; DB 14; Length 1191;
Best Local Similarity 48.5%; Pred. No. 13;
Matches 133; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```

```

QY 675 GTAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTCCATCTCTAGCAACATA 734
Db 858 GTGATATTCTTGACATTAGCCGATAAAGTCTGCAAGACAGAGACTGAAGCGTGCACG 799
QY 735 ACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCACAGGAAACGC 794
Db 798 CCACAGATATCAAGCTTTTAAAGACGAGATGGAGAGAGATGTAAGAAGACGATATTC 739
QY 795 CTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAGATAGAGAAG 854
Db 738 CCACTATTGTACATTCTACAGATCAAAAACCATCTCTGATAAAGTGTATTGAAGAAAGAAG 679
QY 855 ATGGAATGCGATTAGGATGATATTTTGCAGGACAGACGCTAGTCGAGCTGTAGAAA 914
Db 678 AGGAAGAAGATGGCGGGAAGAGAGAGAGAGAAATCAATAGAGAAGAGAAAGTA 619
QY 915 ATGCTAAATCTAATAACAGTATAAGCAACATAGA 948
Db 618 TTGGTAAGAGTGATACGTCTGTACCATATAAAGA 585

RESULT 6
BX353273/c
LOCUS
DEFINITION BX353273 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC003YH11 5-PRIME, mRNA sequence.
ACCESSION BX353273
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC003CD06Q21.
FEATURES
source
1..533
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC003YH11"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dr)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 2.2%; Score 48.2; DB 13; Length 533;
Best Local Similarity 41.4%; Pred. No. 16;
Matches 48; Conservative 35; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGGCTCGCATTAATTATTAGTGAGGCTTTTATTTATTTT 60
Db 431 AAAAAATATWWWWWWWATTTTWWTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 372
QY 61 ATAAATAAATAAAGATTTTATTATTATTTTATGTTTATGTTTATGTTTATGTT 116
Db 371 WWWWWWWWWWWWWHHHHHTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTCNAC 316

RESULT 7

```

```

CK097864
LOCUS
DEFINITION UB64CPC02.3R Populus active cambium cDNA library Populus tremula
cDNA clone UB64CPC02 3', mRNA sequence.
ACCESSION CK097864
VERSION CK097864.1 GI:38582189
KEYWORDS EST.
SOURCE
ORGANISM Populus tremula
Populus tremula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE 1 (bases 1 to 872)
AUTHORS Sterky,F., Bhalerao,R.R., Unneberg,P., Segerman,B., Nilsson,P.,
Brunner,A.M., Campaa,L., Jonsson-Lindvall,J., Tandré,K.,
Strauss,S.H., Sundberg,B., Gustafsson,P., Uhlen,M., Bhalerao,R.P.,
Nilsson,O., Sandberg,G., Karlsson,J., Lundberg,J. and Jansson,S.
A Populus EST resource for functional genomics
Unpublished (2003)
JOURNAL
COMMENT Other_ESTs: UB64CPC02
Contact: Bo Segerman
Umea Plant Science Center, Department of Plant Physiology
Umea University
901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: bo.segerman@plantphys.umu.se.
FEATURES
source
1..872
/organism="Populus tremula"
/mol_type="mRNA"
/db_xref="taxon:113636"
/clone="UB64CPC02"
/tissue_type="Active cambium"
/clone_lib="Populus active cambium cDNA library"

ORIGIN
Query Match 2.2%; Score 48; DB 14; Length 872;
Best Local Similarity 50.4%; Pred. No. 17;
Matches 11; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 1261 TGATTCTCAAGCTGCCCAACAGGAGCTCGCAGCACAGTAGACGAAAGCGCTGG 1320
Db 513 TGCTGCTGCTGCTGCTCAATTGTTGGAAGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGC 572
QY 1321 AGATGACAGTGTCTCTCAGCGGCTGGCAGATGCTCAGAAAGCTTTAGAAAGGCGCTTAG 1380
Db 573 TGCTGCAATTGAGGTAGCTGCTGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGCTGC 632
QY 1381 TAAAGCTGGGCAACACAGGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1440
Db 633 TGTGGCTGCTGCTGTTGTAAGCTGTGAGAGCTGCTGCTGCTGCTGCTGCTGTTGAGC 692
QY 1441 TGTGAGCGCAGGAGTTCCTCCCGCTGCAGCAAGTCTTATAGGTCATCTCTA 1492
Db 693 TGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744

RESULT 8
BH714761
LOCUS
DEFINITION BOMGV71R BO 2.3 KB Brassica oleracea genomic clone BOMGV71,
genomic survey sequence.
ACCESSION BH714761
VERSION BH714761.1 GI:18808752
KEYWORDS GSS.
SOURCE
ORGANISM Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 510)
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.

```


aps 0;
TGCT 1333
|||
AGCT 106
GCAA 1393
|
AGGA 166
AGGA 1453
|||
TGCA 226
C 1510
|
C 283

ostomi;

```

and cdNA
me end
t I and
SPORT 6
aps 0;
TTAT 87
|||

```


Db 103 GCTGCTGCGC 93

RESULT 15

BX403499
LOCUS BX403499 Homo sapiens PLACENTA mRNA linear EST 13-MAY-2003
DEFINITION 3-PRIME mRNA sequence.
ACCESSION BX403499
VERSION BX403499.1 GI:30635017
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 427)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CLOBAA002ZH05FP1.

FEATURES

Location/Qualifiers

1..427

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CLOBAA002ZH05"

/tissue_type="PLACENTA"

/clone_lib="Homo sapiens PLACENTA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 2.1%; Score 46; DB 13; Length 427;
Best Local Similarity 31.5%; Pred. No. 46;
Matches 34; Conservative 47; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATAAATCTTTAAACAGCGCTCGATTATATTAGTCAGAGCTTTTATTTTATTTT 60

Db 64 WWWWWWWWWTT 123

QY 61 ATATAAACTAAAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 108

Db 124 WTWTATTAATWAATWATWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 171

Search completed: March 24, 2004, 05:56:05
Job time : 5664.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:24:24 ; Search time 56.3737 Seconds
(without alignments)
3262.843 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MVNPIGPGIDERTPPAD.....SLFIQVLNIGSLGYLIQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	3 AAY71954	Aay71954 Chlamydia
2	3187	100.0	651	4 AAU38899	Aau38899 C. pneumo
3	3187	100.0	651	5 ABB94277	Abb94277 Chlamydia
4	3187	100.0	651	5 ABB90532	Abb90532 Chlamydia
5	3187	100.0	651	7 ADD42665	Add42665 Chlamydia
6	2849	89.4	583	3 AAY71955	Aay71955 5'-trunca
7	2208	69.3	490	3 AAY71957	Aay71957 C. pneumo
8	2204	69.2	452	3 AAY71956	Aay71956 3'-trunca
9	2182	68.5	478	2 AAY35358	Aay35358 Chlamydia
10	1000	31.4	212	2 AAY35357	Aay35357 Chlamydia
11	899.5	28.2	647	5 ABB94283	Abb94283 Chlamydia
12	899.5	28.2	647	7 ADD43866	Add43866 Chlamydia
13	894	28.1	660	4 AAU38921	Aau38921 C. tracho
14	894	28.1	660	7 ADD42726	Add42726 Chlamydia
15	671.5	21.1	361	3 AAB13695	Aab13695 Chlamydia
16	671.5	21.1	361	4 AAG83263	Aag83263 Protein e
17	671.5	21.1	361	5 ABB94234	Abb94234 Chlamydia
18	604.5	19.0	331	2 AAY37572	Aay37572 Chlamydia
19	289.5	9.1	350	2 AAY37571	Aay37571 Chlamydia
20	251	7.9	1463	5 AAE20110	Aae20110 Lactobaci
21	235.5	7.4	1178	6 ABB26182	Abb26182 Aspergill
22	234	7.3	1831	6 ABU43109	Abu43109 Protein e
23	234	7.3	1870	6 ABJ19019	Abj19019 Pathogen
24	233	7.3	1822	2 AAR27745	Aar27745 Extracell
25	231.5	7.3	2478	4 AAU37374	Aau37374 Staphyloc

26	231.5	7.3	2478	4 AAU34320	Aau34320 Staphyloc
27	231.5	7.3	2478	6 ABJ19002	Abj19002 Pathogen
28	231.5	7.3	2478	6 ABM71899	Abm71899 Staphyloc
29	223.5	7.0	2481	6 ABU15838	Abu15838 Protein e
30	223.5	7.0	2481	7 ABR62804	AbR62804 Methicill
31	221.5	7.0	971	7 ADC00627	Adc00627 Enterohae
32	218	6.8	603	6 ABU44080	Abu44080 Protein e
33	217.5	6.8	6713	6 ABU15887	Abu15887 Protein e
34	217	6.8	5795	4 AAU37017	Aau37017 Staphyloc
35	215.5	6.8	2016	6 AAE36891	Aae36891 Plectreur
36	214.5	6.7	2137	5 ABP39618	Abp39618 Staphyloc
37	213	6.7	2434	4 AAU34339	Aau34339 Staphyloc
38	213	6.7	6281	4 AAU37403	Aau37403 Staphyloc
39	213	6.7	10498	6 ABU19119	Abu19119 Pathogen
40	210	6.6	1122	6 ABU14859	Abu14859 Protein e
41	210	6.6	1123	6 ABU42657	Abu42657 Protein e
42	210	6.6	2271	6 ABU16000	Abu16000 Protein e
43	210	6.6	10182	5 ABP38314	Abp38314 Staphyloc
44	208.5	6.5	4776	6 ABU02253	Abu02253 S. pneumo
45	208.5	6.5	4776	6 ABU45754	Abu45754 Protein e

ALIGNMENTS

RESULT 1
AAY71954
ID AAY71954 standard; protein; 651 AA.
XX
AC AAY71954;
XX
XX 12-SEP-2003 (revised)
DT 26-MAR-2001 (first entry)
XX
DE Chlamydia pneumoniae 76 kDa full-length protein.
XX

76 kDa protein; bactericidal; diagnosis; prevention; treatment;
KW pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation.
XX

OS Chlamydophila pneumoniae.
XX

PN WC000066739-A2.
XX

PD 09-NOV-2000.
XX

PF 03-MAY-2000; 2000WO-CA000511.
XX

PR 03-MAY-1999; 99US-0132270P.
XX

PR 30-JUN-1999; 99US-0141276P.
XX

PA (AVET) AVENTIS PASTEUR LTD.
XX

PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX

WPI: 2000-687542/57.
XX

N-PSDB; AAD02063.
XX

Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.

Claim 16a; Fig 1; 112pp; English.

The present sequence is Chlamydia pneumoniae full-length 76 kDa protein.
C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
tract disease, bronchitis, sinusitis and acute respiratory disease such
as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
auscultation). C. pneumoniae sequence is also used as vaccines for
immunising humans against diseases caused by C. pneumoniae. (Updated on
12-SEP-2003 to standardise OS field)

XX

app. can

SQ Sequence 651 AA;

Query Match 100.0%; Score 3187; DB 3; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.6e-194;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI 60
 DB 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI 60

QY 61 LRSVAVNALSADKLGIASSNSSTSRSDVDSITTATPTPPPTFDYKTKQAQAYDT 120
 DB 61 LRSVAVNALSADKLGIASSNSSTSRSDVDSITTATPTPPPTFDYKTKQAQAYDT 120

QY 121 IFTSTSLADIQAALVSLQDAVTVNIKDTAETETAIAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTVNIKDTAETETAIAAEWETKNADAVKVGAIITELAKY 180

QY 181 ASDNQAILDSLGLTSFDLLQAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIASQSLVD 240
 DB 181 ASDNQAILDSLGLTSFDLLQAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIASQSLVD 240

QY 241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300

QY 301 KKFPDSPILQBAEQMVIAEQKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPDSPILQBAEQMVIAEQKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360

QY 361 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQAAKAAAGDSDSAAALADAK 420
 DB 361 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQAAKAAAGDSDSAAALADAK 420

QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT 480
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT 480

QY 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSPARTEARGPEKTDQALARV 540
 DB 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSPARTEARGPEKTDQALARV 540

QY 541 ISGNSRTLGDVYSQVSALOSVWQIIOSNPOANNEEIRQKLTSAVTKPPQFGYPYVOLSND 600
 DB 541 ISGNSRTLGDVYSQVSALOSVWQIIOSNPOANNEEIRQKLTSAVTKPPQFGYPYVOLSND 600

QY 601 STQKFIKLESLFAEGSRTAABEIKALSPETNSLFIQOVLVNIIGSLYSGYLQ 651
 DB 601 STQKFIKLESLFAEGSRTAABEIKALSPETNSLFIQOVLVNIIGSLYSGYLQ 651

RESULT 2

AAU38899
 ID AAU38899 standard; protein; 651 AA.

XX AC AAU38899;

XX AC AAU38899;

DT 11-SEP-2003 (revised)

DT 16-JAN-2002 (first entry)

XX DE

DE C. pneumoniae CT622 homologue CPn0728.

XX DE

KW Chlamydia; sexually transmitted disease; PID; antibacterial;

KW pelvic inflammatory disease; antigen; trachoma; gynecological;

KW acute respiratory tract infection; atherosclerosis; male infertility;

KW coronary heart disease.

XX OS

OS Chlamydothila pneumoniae.

XX OS

PN W0200181379.A2.

XX OS

PD 01-NOV-2001

XX OS

PF 23-APR-2001; 2001WO-US013081.
 XX PR 21-APR-2000; 2000US-0198853P.
 PR 20-JUL-2000; 2000US-0219752P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Bhatia A, Probst P, Stromberg EJ;
 XX DR MPI; 2001-616771/71.
 XX DR N-PSDB; AAS57031.
 XX PT New polynucleotide for treating Chlamydia infections encodes a
 PT polynucleotide **containing** an immunogenic portion of a Chlamydia antigen.
 XX PS Disclosure; Page 161-162; 208pp; English.
 XX CC The invention relates to isolated polynucleotide encoding at least a
 CC partial Chlamydia protein which is an antigenic fragment, or the
 CC complements, fragments, homologues and variants, and antibodies raised
 CC against the antigenic proteins (or fragments). The nucleic acids,
 CC proteins and antibodies are used to diagnose and treat Chlamydia
 CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
 CC disease (PID), acute respiratory tract infection, trachoma,
 CC atherosclerosis and coronary heart disease) in a patient, and in the
 CC treatment of male infertility. The compounds of the invention are also
 CC useful for detecting the presence of Chlamydia in a patient, and
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-
 CC 2003 to standardise OS field)
 XX CC
 SQ Sequence 651 AA;
 Query Match 100.0%; Score 3187; DB 4; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.6e-194;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI 60
 DB 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAQAQIAGAAKPKSKTDSVERWSI 60

QY 61 LRSVAVNALSADKLGIASSNSSTSRSDVDSITTATPTPPPTFDYKTKQAQAYDT 120
 DB 61 LRSVAVNALSADKLGIASSNSSTSRSDVDSITTATPTPPPTFDYKTKQAQAYDT 120

QY 121 IFTSTSLADIQAALVSLQDAVTVNIKDTAETETAIAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTVNIKDTAETETAIAAEWETKNADAVKVGAIITELAKY 180

QY 181 ASDNQAILDSLGLTSFDLLQAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIASQSLVD 240
 DB 181 ASDNQAILDSLGLTSFDLLQAALLOSVAANNKAAELLKEMQDNVPVPGKTPAIASQSLVD 240

QY 241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300

QY 301 KKFPDSPILQBAEQMVIAEQKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPDSPILQBAEQMVIAEQKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360

QY 361 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQAAKAAAGDSDSAAALADAK 420
 DB 361 DDAENETASILMSGFRQMIHMENTENPDQAQOELAAQAAKAAAGDSDSAAALADAK 420

QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT 480
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASIGSSVKQLYKTSKSGDYKT 480

QY 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSPARTEARGPEKTDQALARV 540
 DB 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSPARTEARGPEKTDQALARV 540

QY 541 ISGNSRTGLDVYSQVSALQSVQVQIIQSNPOANNEIROKLTSAVTKPQFGYPVQLSND 600
 Db 541 ISGNSRTGLDVYSQVSALQSVQVQIIQSNPOANNEIROKLTSAVTKPQFGYPVQLSND 600
 QY 601 STQKFIKLSLFAEGSRTAAEIKALSPEFNSLFIQQVLVNIQSLYSYLQ 651
 Db 601 STQKFIKLSLFAEGSRTAAEIKALSPEFNSLFIQQVLVNIQSLYSYLQ 651

RESULT 3
 ABB94277 ID ABB94277 standard; protein; 651 AA.
 XX
 AC ABB94277;
 XX
 DT 29-AUG-2003 (revised)
 DT 05-JUN-2002 (first entry)
 XX
 XX Chlamydia pneumoniae protein sequence SEQ ID NO:405.

XX Chlamydia pneumoniae; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW antibacterial; immunostimulant; immune response;
 KW Chlamydia-specific T-cell response.

OS Chlamydia pneumoniae.

XX WO200208267-A2.
 XX 31-JAN-2002.
 XX 20-JUL-2001; 2001WO-US023121.
 XX 20-JUL-2000; 2000US-00620412.
 PR 23-APR-2001; 2001US-00841132.
 XX
 XX (CORI-) CORIXA CORP.

PI Fling SP, Skeiky YAW, Probst P, Bhatia A;
 XX WPI; 2002-179901/23.

XX Novel compositions comprising Chlamydia CapI protein and its use in the
 PT treatment of Chlamydia infection.

XX Disclosure; Page 370-372; 537pp; English.

XX The present invention describes compositions comprising a Chlamydia CapI
 CC protein and methods for the diagnosis and therapy of Chlamydia infection.
 CC Chlamydia DNA and protein sequences from the present invention can have
 CC antibacterial and immunostimulant activities, and can be used in
 CC vaccines. Compounds from the present invention can be used for eliciting
 CC an immune response, specifically stimulating a Chlamydia-specific T-cell
 CC response or inhibiting the development of a Chlamydia infection in an
 CC animal. Methods from the present invention can be used for detecting the
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia
 CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
 CC sequences used in the exemplification of the present invention. (Updated
 CC on 29-AUG-2003 to standardise OS field)

XX Sequence 651 AA;

Query Match 100.0%; Score 3187; DB 5; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.6e-194;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNPIGPGFIDETERTPADLSAQGLEASANKSAEQRIAGAEAKPKESKTSVERWSI 60
 Db 1 MVNPIGPGFIDETERTPADLSAQGLEASANKSAEQRIAGAEAKPKESKTSVERWSI 60
 QY 61 LRSVNLMSLADKLGIASSNSSSTSRSDVDSSTATATPPPPPTFDYKTAQAYDT 120
 Db 61 LRSVNLMSLADKLGIASSNSSSTSRSDVDSSTATATPPPPPTFDYKTAQAYDT 120

QY 121 IFTSTSLADIQAALVSLQDAVTNIKOTAAATDEETAAIAEWETKNADAVKVGQAQITELAKY 180
 Db 121 IFTSTSLADIQAALVSLQDAVTNIKOTAAATDEETAAIAEWETKNADAVKVGQAQITELAKY 180
 QY 181 ASDNQAILDSLGKLTSPDLQAALLQSVANNKAAELLKEMQDNPVVPKTPPAQSLVD 240
 Db 181 ASDNQAILDSLGKLTSPDLQAALLQSVANNKAAELLKEMQDNPVVPKTPPAQSLVD 240
 QY 241 QTDATATQIEKDGNAIIRDYFAGQNASGAVENAKSNNSISINIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKDGNAIIRDYFAGQNASGAVENAKSNNSISINIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFDPSPILQEAQOMVIOAEKDLKNIKPADGSDVPNFGTTVGGSKQGGSSIGSRVSMML 360
 Db 301 KKFDPSPILQEAQOMVIOAEKDLKNIKPADGSDVPNFGTTVGGSKQGGSSIGSRVSMML 360
 QY 361 DDAENETASILMSGFROMIHMFNTENPDSQAACQELAAQAPAAKAAAGDSDSAAALADAOK 420
 Db 361 DDAENETASILMSGFROMIHMFNTENPDSQAACQELAAQAPAAKAAAGDSDSAAALADAOK 420
 QY 421 ALEAALGKAGQQQGIILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 Db 421 ALEAALGKAGQQQGIILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 QY 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTALTTSVPRAETARGPEKTDQALRV 540
 Db 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTALTTSVPRAETARGPEKTDQALRV 540
 QY 541 ISGNSRTGLDVYSQVSALQSVQVQIIQSNPOANNEIROKLTSAVTKPQFGYPVQLSND 600
 Db 541 ISGNSRTGLDVYSQVSALQSVQVQIIQSNPOANNEIROKLTSAVTKPQFGYPVQLSND 600
 QY 601 STQKFIKLSLFAEGSRTAAEIKALSPEFNSLFIQQVLVNIQSLYSYLQ 651
 Db 601 STQKFIKLSLFAEGSRTAAEIKALSPEFNSLFIQQVLVNIQSLYSYLQ 651

RESULT 4

ABB90532 ID ABB90532 standard; protein; 651 AA.

XX AC ABB90532;

XX DT 29-AUG-2003 (revised)

XX DT 29-JUL-2002 (first entry)

XX Chlamydia pneumoniae cp7033 protein, SEQ ID NO:13.

XX Chlamydia pneumoniae; antigen; immunogen; vaccine; diagnosis;
 KW human respiratory disease; cardiovascular disease; atherosclerosis;
 KW coronary artery disease; carotid artery stenosis; myocardial infarction;
 KW cerebrovascular disease; aortic aneurysm; claudication; stroke;
 strain CWL029.

OS Chlamydia pneumoniae.

XX WO200202606-A2.

XX 10-JAN-2002.

XX 03-JUL-2001; 2001WO-IB001445.

XX 03-JUL-2000; 2000GB-00016363.

XX 11-JUL-2000; 2000GB-00017047.

XX 21-JUL-2000; 2000GB-00017983.

XX 07-AUG-2000; 2000GB-00019368.

XX 18-AUG-2000; 2000GB-00020440.

XX 14-SEP-2000; 2000GB-00022583.

XX 10-NOV-2000; 2000GB-00027549.

XX 22-DEC-2000; 2000GB-00031706.

XX (CHIR-) CHIRON SPA.

XX Ratti G, Grandi G;
XX WPI: 2002-154726/20.
XX N-PSDB; ABU91190.
XX Novel Chlamydia pneumoniae protein useful in the manufacture of a
XX medicament for treatment or prevention of infection due to Chlamydia,
XX preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX Claim 1; Page 47; 364pp; English.
XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
XX pneumoniae (strain CML029), and ABU91184-ABU91373 represent DNA encoding
XX them. The proteins are predicted to be immunogenic and may therefore be
XX useful in vaccine production and for diagnostic purposes. Chlamydia
XX pneumoniae is a common cause of respiratory disease in humans, and is
XX also involved in the development of cardiovascular diseases such as
XX atherosclerosis, coronary artery disease, carotid artery stenosis,
XX myocardial infarction, cerebrovascular disease, aortic aneurysm,
XX claudication and stroke. The proteins and nucleic acids of the invention
XX may be used in vaccines and pharmaceutical compositions for the
XX prevention or treatment of chlamydial infections, particularly Chlamydia
XX pneumoniae infections. The proteins may also be used in the detection of
XX Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
XX DNA probe assay or blotting techniques for determining Chlamydia
XX pneumoniae gene expression. The present sequence represents a
XX specifically claimed Chlamydia pneumoniae protein of the invention.
XX (Updated on 29-AUG-2003 to standardise OS field)
XX Sequence 651 AA;
Query Match 100.0%; Score 3187; DB 5; Length 651;
Best Local Similarity 100.0%; Pred No. 4.6e-194; Indels 0; Gaps 0;
Matches 651; Conservative 0; Mismatches 0;
QY 1 MWNPTGGPIDETERTPPADLSAQGLESAANKSABEAQRIAGAEKPKESKTSVERWSI 60
DB 1 MWNPTGGPIDETERTPPADLSAQGLESAANKSABEAQRIAGAEKPKESKTSVERWSI 60
QY 61 LRSVNAWMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
DB 61 LRSVNAWMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
QY 121 IFTSTSLADIQAALVSLQDVAVNIKDTAETAEWETKNADAVKVAQITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDVAVNIKDTAETAEWETKNADAVKVAQITELAKY 180
QY 181 ASDNQAILDSIGKLTSLFLLQALILQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
DB 181 ASDNQAILDSIGKLTSLFLLQALILQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
QY 241 QTDATATQIEKGNRAIRDAYFAGQNASGAVENAKNSNISIDSAAIAATAKTOIAEAQ 300
DB 241 QTDATATQIEKGNRAIRDAYFAGQNASGAVENAKNSNISIDSAAIAATAKTOIAEAQ 300
QY 301 KKFDPSPILQEAQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQGGSSIGSRVSMML 360
DB 301 KKFDPSPILQEAQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQGGSSIGSRVSMML 360
QY 361 DDAENETASILMSGFRQIMHNTENPDPSQAQQLAAQAAKAGDSDSAAALADQK 420
DB 361 DDAENETASILMSGFRQIMHNTENPDPSQAQQLAAQAAKAGDSDSAAALADQK 420
QY 421 ALEAALGAGQOQGLNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSGDYKT 480
DB 421 ALEAALGAGQOQGLNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSGDYKT 480
QY 481 QISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSPVPRTEARGPKTKDQALRV 540
DB 481 QISAGYDAYKSINDAYGRANDATEDVDVINNVSTPALTRSPVPRTEARGPKTKDQALRV 540
QY 541 ISGNSRTIGDVYSQVSAALQSVNQIIQSNPQANNEIROKLTSAVTKPPQFGYPYQLSND 600

DB 541 ISGNSRTIGDVYSQVSAALQSVNQIIQSNPQANNEIROKLTSAVTKPPQFGYPYQLSND 600
QY 601 STQKFIKLESIFABGSRTAAEIKALSFTETNSLFTQQVLVNIIGSLYSGLYQ 651
DB 601 STQKFIKLESIFABGSRTAAEIKALSFTETNSLFTQQVLVNIIGSLYSGLYQ 651
RESULT 5
ADD42665
ID ADD42665 standard; protein; 651 AA.
XX AC ADD42665;
XX DT 15-JAN-2004 (first entry)
XX DE Chlamydia pneumoniae antigen protein SEQ ID NO:78.
XX KW Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
XX KW antiinfertility; cardiant; antiarteriosclerotic; ophthalmological;
KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
KW tubal obstruction; infertility; male infertility; ocular infection;
KW blindness; acute respiratory tract infection; atherosclerosis;
XX KW coronary heart disease.
OS Chlamydia pneumoniae.
XX WO2003041560-A2.
XX PD 22-MAY-2003.
XX PF 05-NOV-2002; 2002WO-US035624.
XX PR 06-NOV-2001; 2001US-00012256;
PR 08-DEC-2001; 2001US-00007693.
PR 15-JUL-2002; 2002US-00197220.
XX PA (CORI) CORIXA CORP.
XX PI Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;
PI Probst P;
XX DR WPI: 2003-441771/41.
XX DR N-PSDB; ADD42625.
XX PT New DNA and proteins comprising a portion of a Chlamydia antigen, useful
PT for diagnosing or treating Chlamydial infections, particularly as
PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
XX inflammatory disease.
PS Disclosure; SEQ ID NO 78; 275pp; English.
XX The present invention describes compounds and methods for diagnosing and
XX treating Chlamydial infection. Chlamydia polynucleotide and protein
XX sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
XX antiarteriosclerotic and ophthalmological activities, and can be used in
XX vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
XX compositions or methods from the present invention can be used for the
XX serodiagnosis or treatment of Chlamydial infections, particularly in
XX humans. The polynucleotides, proteins or compositions are particularly
XX useful for stimulating an immune response in a patient, or for
XX stimulating and/or expanding T cells specific for a Chlamydia protein.
XX Specifically, the polynucleotides, proteins or compositions are useful as
XX vaccines for treating or preventing Chlamydial infections including
XX pelvic inflammatory disease (which results in tubal obstruction and
XX infertility in women), male infertility, ocular infection (which may
XX cause blindness), acute respiratory tract infections, atherosclerosis, or
XX coronary heart disease. The present sequence is used in the
XX exemplification of the present invention.
XX Sequence 651 AA;
Query Match 100.0%; Score 3187; DB 7; Length 651;

RESULT 7
 AAY71957
 ID AAY71957 standard; protein; 490 AA.
 AC AAY71957;
 XX
 XX
 DT 26-MAR-2001 (first entry)
 XX
 XX C. pneumoniae 76 kDa protein truncation mutant fusion protein.
 XX
 XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; fusion protein; truncation mutant;
 XX mutin.
 XX
 XX Chlamydia pneumoniae.
 OS Synthetic.
 XX
 XX
 FH Key Location/Qualifiers
 FT 453..490
 FT Region
 FT /note= "this part of the sequence is unrelated to the C.
 FT pneumoniae 76 kDa protein"
 XX
 XX WO200066739-A2.
 PN
 XX
 XX 09-NOV-2000.
 PD
 XX
 XX 03-MAY-2000; 2000WO-CA000511.
 PF
 XX
 XX 03-MAY-1999; 99US-0132270P.
 PR 30-JUN-1999; 99US-0141276P.
 XX
 XX (AVET) AVENTIS PASTEUR LTD.
 PA
 XX
 XX Murdin AD, Oomen RP, Wang J, Dunn P;
 PI WPI; 2000-687542/67.
 XX N-PSDB; AAD02066.
 DR
 XX
 XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 PT for vaccinating against Chlamydia infections.
 PT
 XX
 XX Claim 33; Fig 3; 112pp; English.
 PS
 XX The present sequence is a fusion protein comprising a truncated Chlamydia
 CC pneumoniae 76 kDa protein and vector-encoded residues. C. pneumoniae 76
 CC kDa protein is used in the diagnosis, prevention and treatment of C.
 CC pneumoniae infections (e.g. pneumonia, upper respiratory tract disease,
 CC bronchitis, sinusitis and acute respiratory disease such as cough, sore
 CC throat, hoarseness, fever; and abnormal chest sounds on auscultation). C.
 CC pneumoniae sequence is also used as vaccines for immunising humans
 CC against diseases caused by C. pneumoniae
 XX
 XX Sequence 490 AA;
 SQ
 Query Match 69.3%; Score 2208; DB 3; Length 490;
 Best Local Similarity 99.8%; Pred. No. 5.1e-132;
 .Matches 453; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAEAKPKESKTSVERWSI 60
 DB 1 MNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAEAKPKESKTSVERWSI 60
 QY 61 LRSVNMALSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDYKTKQAQYDT 120
 DB 61 LRSVNMALSLADKLGIASSNSSSTSRADVDSTTATPTPPPTFDYKTKQAQYDT 120
 QY 121 IFTTSLADIQAALVSLQDAVNIKDTAATDBETALAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTTSLADIQAALVSLQDAVNIKDTAATDBETALAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSLGLTSLFDLLQALQLQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240

Db 181 ASDNQAILDSLGLTSLFDLLQALQLQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 QY 241 QTDATAQTQTEKDNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAATATAKTQIAEAQ 300
 Db 241 QTDATAQTQTEKDNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAATATAKTQIAEAQ 300
 QY 301 KKFPPSPILQEAQWVIAEKDLKNIKPADGSDVNPVPGTTVGGSKQGGSSIGSIRVSMML 360
 Db 301 KKFPPSPILQEAQWVIAEKDLKNIKPADGSDVNPVPGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMHFNTENPDQAAQQLAQAARAKAAGDSDAAALADAKQ 420
 Db 361 DDAENETASILMSGFRQMHFNTENPDQAAQQLAQAARAKAAGDSDAAALADAKQ 420
 QY 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVPP 454
 Db 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVLP 454
 RESULT 8
 AAY71956
 ID AAY71956 standard; protein; 452 AA.
 XX
 AC AAY71956;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE 3'-truncated Chlamydia pneumoniae 76 kDa protein.
 XX
 KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; truncation mutant; mutin.
 XX
 OS Chlamydia pneumoniae.
 OS Synthetic.
 XX
 XX WO200066739-A2.
 PN
 XX
 XX 09-NOV-2000.
 PD
 XX
 XX 03-MAY-2000; 2000WO-CA000511.
 PF
 XX
 XX 03-MAY-1999; 99US-0132270P.
 PR 30-JUN-1999; 99US-0141276P.
 XX
 XX (AVET) AVENTIS PASTEUR LTD.
 PA
 XX
 XX Murdin AD, Oomen RP, Wang J, Dunn P;
 PI WPI; 2000-687542/67.
 XX N-PSDB; AAD02065.
 DR
 XX
 XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 PT for vaccinating against Chlamydia infections.
 PT
 XX
 XX Claim 16c; Page 104-106; 112pp; English.
 PS
 XX The present sequence is 3'-truncated Chlamydia pneumoniae 76 kDa protein.
 CC pneumoniae 76 kDa protein is used in the diagnosis, prevention and
 CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
 CC tract disease, bronchitis, sinusitis and acute respiratory disease such
 CC as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
 CC auscultation). C. pneumoniae sequence is also used as vaccines for
 CC immunising humans against diseases caused by C. pneumoniae
 XX
 XX Sequence 452 AA;
 SQ
 Query Match 69.2%; Score 2204; DB 3; Length 452;
 Best Local Similarity 100.0%; Pred. No. 8.3e-132;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PA (GEST) GENSET.
 XX Griffais R;
 PI WPI; 1999-357842/30.
 XX Genome sequence of Chlamydia pneumoniae.
 XX PT
 XX PS Page 1155; Disclosure; 1912pp; English.
 XX CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAY91990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -Oct-2003 to standardise OS field)
 XX SQ Sequence 212 AA;
 Query Match 31.4%; Score 1000; DB 2; Length 212;
 Best Local Similarity 99.5%; Pred. No. 9.7e-56;
 Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 448 VSAGVPPAAASIGSGSVKQLYTSTSGDYKQTSAGVDAYKSNIDAYGRARNDTRDV 507
 Db |||||
 QY 9 VSAGVPPAAASIGSGSVNQLYTSKSTSGDYKQTSAGVDAYKSNIDAYGRARNDTRDV 68
 Db |||||
 QY 508 INNVSTPALTRSPVPRARTFARGPEKTDQALRVISGNSRTLDGVYQVSVLSQVMOILOS 567
 Db |||||
 QY 69 INNVSTPALTRSPVPRARTFARGPEKTDQALRVISGNSRTLDGVYQVSVLSQVMOILOS 128
 Db |||||
 QY 568 NPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFTAKLESIFABGSRTPAAIRKALS 627
 Db |||||
 QY 129 NPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFTAKLESIFABGSRTPAAIRKALS 188
 Db |||||
 QY 628 FETNSLFTQOVLNIGSLYSGLQ 651
 Db |||||
 QY 189 FETNSLFTQOVLNIGSLYSGLQ 212
 Db |||||
 RESULT 11
 ABB94283
 ID ABB94283 standard; protein; 647 AA.
 AC ABB94283;
 XX 05-JUN-2002 (first entry)
 DE Chlamydia trachomatis protein sequence SEQ ID NO:436.
 XX Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
 KW antibacterial; immunostimulant; immune response;
 KW Chlamydia-specific T-cell response.
 XX Chlamydia trachomatis.
 OS Chlamydia trachomatis.
 XX WO200208267-A2.
 FN 31-JAN-2002.
 PD 20-JUL-2001; 2001WO-US023121.
 PF 20-JUL-2000; 2000US-00620412.
 PR 23-APR-2001; 2001US-00841132.
 XX (CORI-) CORIXA CORP.
 PA Fling SP, Skeiky YAW, Probst P, Bhatia A;
 PI

XX WPI; 2002-179901/23.
 DR Novel compositions comprising Chlamydia CapI protein and its use in the
 XX treatment of Chlamydia infection.
 PT Claim 34; Page 391-393; 537pp; English.
 XX The present invention describes compositions comprising a Chlamydia CapI
 CC protein and methods for the diagnosis and therapy of Chlamydia infection.
 CC Chlamydia DNA and protein sequences from the present invention can have
 CC antibacterial and immunostimulant activities, and can be used in
 CC vaccines. Compounds from the present invention can be used for eliciting
 CC an immune response, specifically stimulating a Chlamydia-specific T-cell
 CC response or inhibiting the development of a Chlamydia infection in an
 CC animal. Methods from the present invention can be used for detecting the
 CC presence of Chlamydia in a patient; to stimulate and/or expand T cells
 CC specific for a Chlamydia protein; and for treatment of a Chlamydia
 CC infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
 CC sequences used in the exemplification of the present invention
 XX SQ Sequence 647 AA;
 Query Match 28.2%; Score 899.5; DB 5; Length 647;
 Best Local Similarity 35.1%; Pred. No. 1e-48;
 Matches 234; Conservative 118; Mismatches 262; Indels 53; Gaps 16;
 QY 2 VNPICGPDIDETTPPADLSAQLGSAANKSABAQRIAGAAKPKSKTDSVERMSIL 61
 Db |||||
 QY 15 MNPIINGQI-----ASNSETKESKSEA-----SPSASSSVSSWGFL 52
 Db |||||
 QY 62 RSANVNLMSLADKLGIASSSSSSSTSR8-ADVDSITTATPTPPPTDDYKQTAIDT 120
 Db |||||
 QY 53 SSXKHALISRD--AILNKNSSPTDSLSEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110
 Db |||||
 QY 121 IFTSTSLADIQAALVSLQDAVNIKDTAATDEE-TAJAAEWETKNDAVKVGAQITELAK 179
 Db |||||
 QY 111 LENATTLAEYETKADLMAALQDMERLAKOKAEVTRKEALQEKQ----EVIDKLNOLVK 166
 QY 180 YASNQAILDSGLKLTSPDLLQALLOSANNVAAELLKEMQNPVVGKTPAIAGSLV 239
 Db |||||
 QY 167 LEKQNTLKETLTITTSADQIPAINSOLEINKNSADQIIKLEQNI---SYEAVLTNAG 223
 Db |||||
 QY 240 DQDATATQIEKDGNAIRDYFAGNAGAVENAKSNNSISNIDSAKAAITAKTQAEAA 299
 Db |||||
 QY 224 EVIKASSEAGHLQALQSLVDAGDQQAALVQANNSPDNIATKLLDAAETKKNEL 283
 QY 300 QKK---FPDSPILQEAQWVIAEKDLKNIKPADGSDVPNPGTTVGSKQGGSSIGSI-- 354
 Db |||||
 QY 284 KQHTGLTDSPLVKKAEQISQAQKDIQEKPS-GSDIPIVGPS-GSAASAGSAGVAGALS 341
 QY 355 -----RVSMILLDDAENETASILMSGFFQIMHFNTEPNPDSQAQOELAAQRAAK---AA 406
 Db |||||
 QY 342 SNNSGRISLLDDVDNEMAAIAMQGFPSMIEQFNWNPATAKELQAMEAQLTAMSDQLVG 401
 QY 407 GDSAAALADAQXALEALCKAGQOQOILNALQIASAAVWSAGVPPAAASISGSSVKQ 466
 Db |||||
 QY 402 ADGELPAEIQAIDALACAL-KQSTDGLATAMQVAFAAKVGSGSAGTAGTQVMNVKQ 460
 QY 467 LYKT--SKSTGSDYKTOISAGYDAYKSNIDAYGRARNDTRDVINNVTSTPALTRSPRAR 524
 Db |||||
 QY 461 LYKTAFTSTSSSSYAAALSDGYSAYKTLNLSYBSRS-GVQSAISQTANPALSRVSRSRG 519
 QY 525 TEAREG-PEKTDQALRVISGNSRTLDGVYQVSVLSQVMOILOSNNPANNHEETPKLTSA 583
 Db |||||
 QY 520 IESQSRADASQRAAETIVRDSQTLGDVYSLRLQVLDLSMTSTIVSNPQVNFEEINQKLTAS 579
 QY 584 VTKPPQFGYPVQLSNDSTQKFTAKLESIFABGSRTPAAIRKALSFTNSLFIQOVLNIG 643
 Db |||||
 QY 580 ISKAPQFGYPAVQNSADSLQKFAAQLEREFDGERSLAESRENAFRKQPAFIQOVLNIA 639
 QY 644 SLYSGYL 650
 Db |||||

Db 640 SLFSGYL 646

RESULT 12

ADD43866

ID ADD43866 standard; protein; 647 AA.

XX

AC ADD43866;

XX

XX 15-JAN-2004 (first entry)

DT

DE Chlamydia trachomatis immunogenic protein, SEQ ID NO 161.

XX

XX Chlamydia trachomatis

XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;

XX gene therapy; antibacterial.

XX

OS Chlamydia trachomatis.

XX

XX WO2003049762-A2.

FN

XX 19-JUN-2003.

PD

PF 12-DEC-2002; 2002WO-IB005761.

XX

XX 12-DEC-2001; 2001GB-00029732.

PR

PR 06-AUG-2002; 2002GB-00018233.

PR

XX 14-AUG-2002; 2002GB-00018924.

XX

PA (CHIR-) CHIRON SPA.

XX

XX Grandi G, Ratti G;

FI

DR WPI; 2003-532882/50.

DR

DR N-PSDB; ADD43867.

XX

XX New immunogenic composition having a protein or encoding nucleic acid,

PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis

PT infection.

XX

XX Claim 6; SEQ ID NO 161; 164pp; English.

XX

CC The invention relates to a novel immunogenic composition comprising a

CC protein or nucleic acid, and an adjuvant, where the protein or nucleic

CC acid comprises any of 131 fully defined amino acid or nucleotide

CC sequences given in the specification, or has 50% or greater sequence

CC identity to it, or their fragments. The protein and/or nucleic acid of

CC the immunogenic composition is useful in the manufacture of a medicament

CC for the treatment or prevention of infection due to Chlamydia

CC trachomatis. The infection is treated or prevented by the medicament

CC eliciting an immune response which is specific to a C. trachomatis

CC elementary body, or for neutralising C. trachomatis elementary bodies,

CC hence the immunogenic composition can be used in creating a vaccine. The

CC immunogenic compositions can also be used for the diagnosis of C.

CC trachomatis infection. The nucleic acids of the immunogenic compositions

CC can be used to treat disorders by gene therapy. The immunogenic

CC compositions have antibacterial activity. This sequence represents one of

CC the 131 C. trachomatis proteins with immunogenic properties of the

CC invention.

XX

SQ Sequence 647 AA;

Query Match 28.2%; Score 899.5; DB 7; Length 647;

Best Local Similarity 35.1%; Pred. No. 1e-48;

Matches 234; Conservative 118; Mismatches 262; Indels 53; Gaps 16;

QY 2 VNPFGPIDTERTPPADLSAQGLEASANKSAQAQRIAGAAKPKESKTDSTVERNSIL 61

Db 15 MNPFGQI-----ASNSETKEKSEA-----SPSASSSVSSWSFL 52

QY 62 RSANVALMSLADKLGIASNSSSSTSRG-ADVDSSTATATPTPPPTPDYKTAQATYDT 120

Db 53 SSAGHALISLRD--ALLNKNSSPTLSQLEASTSTSVTRVAARDYNEAKSNFDTAKSG 110

QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEE-TATAAEWETKNADAVKYGAOITELAK 179

Db 111 LENATTIAEYETKADLMALQDNERLAKOKAEVTRIKELQEKQ-----EVIDKLNLQVVK 166

QY 180 YASNQAILDSLGLKLTSLFDLLQALLQSVANNKKAELKEMQDNPVVPVPGKTPAIAQSLV 239

Db 167 LEKQNTLKETLTITTTSDAQIPAINSOLEINKNSADQIINKLEGQNI---SYEAVLTNAG 223

QY 240 DQDQATATQTEKOGNAIRDAYFAQONASGAVENAKSNNSINSDSAKAAIATATQIAEA 299

Db 224 EVIKASSEAGIKLQALQSIQVADGDSQAAVLAQQQNSPDNIAATKLLIDAAETKYNEL 283

QY 300 QKK---FPDPSILOEABQWVIAEKDLKNIKPADGSDVPENPGTTVGGSKQGGSIGSI-- 354

Db 284 QKEHTGLTDSPLVKABEQISQAQKDIQEKPS-GSDIPIVGPS-GSAAAGSAGVAGLKS 341

QY 355 -----RVSMILLDDAENETASILMSGFQIMFMFNTENPDQAAQOELAAQARA--AA 406

Db 342 SNNSGRISLLDDVDNEMAAIAMQGFERSMIEQFNVNPNFATAKELQAMEAQLTAMSDQLVG 401

QY 407 GDSAAALADAQKALEALGKAGQOQOQILNALGOIASAAVVSAGVPPAAASSIGSVKQ 466

Db 402 ADGELPAEIQAIKDALAQAL-KQSTDTGLATAMQGVAFAAKVGGSAGTAGTVQMVVKQ 460

QY 467 LYKT--SKSTGSDYKTIQISAGYDAYKSDINDAYGRANDATRDVINNVSTPALTRSVPRAR 524

Db 461 LYKTAFSTSSSYAAALSDGYSAYKTLNLSYSESRG-GVQSALSQTANFALSVSRSRG 519

QY 525 TEARG-PEKTDQALARVISGNRTGLGVYSQVSLQSMQIIOGSPQANNEBTRKLTSA 583

Db 520 IESQGRSADASQRAAETIVRDSQTLGSDVYSRLQVLDLSLMTIVSNPQVNOEEMOKLTAS 579

QY 584 VTKEPQFGYPVOLSNDSQTKFTAKLSLFAEGSRTAETIKALSFTNSLFIOQVLANIG 643

Db 580 ISKAPQEGYPAVQNSADSLQKFAQLERFVDGERSLAESRENAPRKQAPFIOQVLANIA 639

QY 644 SLXSGYL 650

Db 640 SLFSGYL 646

RESULT 13

AAU38921

ID AAU38921 standard; protein; 660 AA.

XX

AC AAU38921;

XX

DT 11-SEP-2003 (revised)

DT 16-JAN-2002 (first entry)

XX

DE C. trachomatis CT875 protein.

XX

XX Chlamydia; sexually transmitted disease; PID; antibacterial;

XX pelvic inflammatory disease; antigen; trachoma; gynecological;

XX acute respiratory tract infection; atherosclerosis; male infertility;

XX coronary heart disease.

XX

OS Chlamydia trachomatis; serovar E.

XX

FN WO200181379-A2.

XX

XX 01-NOV-2001.

PD

XX 23-APR-2001; 2001WO-US013081.

PF

XX 21-APR-2000; 2000US-0198853P.

PR

XX 20-JUL-2000; 2000US-0219752P.

PR

XX (CORI-) CORIXA CORP.

PA

XX Bhatia A, Probst P, Stromberg EJ;

PI

XX WPI; 2001-616771/71.

DR

DR	N-PSDB; AAS56395.
XX	
XX	
PT	New polynucleotide for treating Chlamydia infections encodes a
PT	polynucleotides containing an immunogenic portion of a Chlamydia antigen.
XX	
PS	Claim 3; Page 205-207; 208pp; English.
XX	
CC	The invention relates to isolated polynucleotide encoding at least a
CC	partial Chlamydia protein which is an antigenic fragment, or the
CC	complements, fragments, homologues and variants, and antibodies raised
CC	against the antigenic proteins (or fragments). The nucleic acids,
CC	proteins and antibodies are used to diagnose and treat Chlamydia
CC	infections (e.g. a sexually transmitted disease, pelvic inflammatory
CC	disease (PID), acute respiratory tract infection, trachoma,
CC	atherosclerosis and coronary heart disease) in a patient, and in the
CC	treatment of male infertility. The compounds of the invention are also
CC	useful for detecting the presence of Chlamydia in a patient, and
CC	stimulating and/or expanding T cells specific for a Chlamydia protein.
CC	The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-
CC	2003 to standardise OS field)
XX	
XX	Sequence 660 AA;
SQ	

```

Query Match      28.1%; Score 894; DB 4; Length 660;
Best Local Similarity 34.6%; Pred.No. 2.4e-48;
Matches 233; Conservative 126; Mismatches 262; Indels 52; Gaps 17;

QY      6  GPGFIDETERTPADLSAQGLEASANKSAEAOQIAGAEAKPKESKTDYSVERNSILRSAY 65
Db      11  GPESVSNQSMNPINGQIASNSETKSYKA-----SEASP--SASSSVSSWFSLSAK 63

QY      66  NALMSIADKLGIASSNSSSTSRs-ADVDSTTATAPTPPTFDYKTOAQYADTIFTS 124
Db      64  NALISLRD--AIINKNSPTDSLSQLSEASTSTSTVTRVAAKYDEAKSNFDTAKSGLENA 121

QY      125  TSLADIQAALVSLQDAVTNIKDTAATD-----EET-----ATAAEWETKNADAVKVGQ 173
Db      122  KTLAEYETKNADLMAALQDMERLANSPPSNHTEEVNNIKALAEAKDITD-----K 173

QY      174  ITELAKYASDNQAILDSIGLKTGFDLLQAALLQSVANNKAAEALLKEMQNVPVPGKTPA 233
Db      174  LNKLVITLQNQKSLTEVLKTTDSADQIPALNSQLEINKNSADQIIKDLERQNI---SYEA 230

QY      234  IAGSLVDQTDATATQTEKGNAINRDYFAGNAGSAGAVENAKSNNSINIDSAAKIATAK 293
Db      231  VLTNAGEVIKASBEAGIKLQCALQSIYDAGDSQAAVLQAQONSPDNIAATKELIDAAE 290

QY      294  TQTAEAQKK---FPDGPILQEAQOMVTAQAEKDKLNKIPADGSDVPNPGTTVGGSKQCGSS 350
Db      291  TKVNELKQEHGTGLTSLPLVKAEEQISOAKDQCEIKPS--GSDIPIVGPS--GSAASAGSA 348

QY      351  IGSII-----RVSMILLDDAENETASTILMSGFQPMHMFNTENPDSSAAQOELAAQAPAA 403
Db      349  AGALKSSNNSGRISLLIILDDVDNEMAAALQGFPSMIEQFNVNPNPATAKEIQAEAOITAM 408

QY      404  K---AAAGDSSAAALADAQKALEAALGKAGQCGGILNALGOIASAAVVSAGVPPAAASSI 460
Db      409  SDQIVGADGELPAEIQAIKDALAQAAL-KQPSADGLATAMQOVAPAAKVGSGSAGTAGTV 467

QY      461  GSSVKQLYKT--SKSTGSDYKTOISAGYDAYKSIINDAYGBARNDATRDVINNVSTPALTR 518
Db      468  QMNVKQLYKTAFTSSTSSSSVAAALSDGYSAYKTLNSLYSESRS--GVQSAISQTPANPALSR 526

QY      519  SVPRARTEARG-PEKTDQALARVIGNSRITLGDVYSQVSAQSALQSVMOIIQSNPQANNEIR 577
Db      527  SVSRSGIESQGRSADASQRAAFTIVRDSQTLGDVYSRLQVLDSLMTSIVSNPQANBEIM 586

QY      578  QKLTSATKPPQFGYPYVQI--SNDSTQKFIAKLBSLFAEGSRTAAEIKALSFTNSLTIQQ 637
Db      587  QKLTSISKAPQFGYPVQNSADSLQKFAQLEREFFVDGERSLAEQENAFRQKPAFIQQ 646

QY      638  VLWNIGSLYSGYL 650
Db      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

Db 647 VLWNTIASLPSGYL 659

RESULT 14
ADD42726
ID ADD42726 standard; protein; 660 AA.
XX AC
XX AC ADD42726;
XX XX
XX 15-JAN-2004 (first entry)
XX XX
XX Chlamydia trachomatis antigen protein SEQ ID NO:139.
XX XX
XX Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
XX XX antiinfertility; cardiant; antiarteriosclerotic; ophthalmological;
XX XX vaccine; gene therapy; immune response; pelvic inflammatory disease;
XX XX tubal obstruction; infertility; male infertility; ocular infection;
XX XX blindness; acute respiratory tract infection; atherosclerosis;
XX XX coronary heart disease.
XX XX
XX Chlamydia trachomatis.
XX OS
XX WO2003041560-A2.
XX PN
XX 22-MAY-2003.
XX PD
XX XX
XX 05-NOV-2002; 2002WO-US035624.
XX PF
XX 06-NOV-2001; 2001US-00012256.
XX PR
XX 05-DEC-2001; 2001US-00007693.
XX PR
XX 15-JUL-2002; 2002US-00157220.
XX PR
XX XX
XX (CORI-) CORIXA CORP.
XX PA
XX XX
XX Bhatia A, Guderian J, Skeiky YAW, Maisonneuve JL, Barth B;
XX PI Probst P;
XX PI
XX XX
XX WPI; 2003-441771/41.
XX DR
XX XX
XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful
XX PT for diagnosing or treating Chlamydial infections, particularly as
XX PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
XX PT inflammatory disease.
XX XX
XX Example 4; SEQ ID NO 139; 275pp; English.
XX PS
XX XX
XX The present invention describes compounds and methods for diagnosing and
XX CC treating Chlamydial infection. Chlamydia polynucleotide and protein
XX CC sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
XX CC antiarteriosclerotic and ophthalmological activities, and can be used in
XX CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
XX CC compositions or methods from the present invention can be used for the
XX CC serodiagnosis or treatment of Chlamydial infections, particularly in
XX CC humans. The polynucleotides, proteins or compositions are particularly
XX CC useful for stimulating an immune response in a patient, or for
XX CC stimulating and/or expanding T cells specific for a Chlamydia protein.
XX CC Specifically, the polynucleotides, proteins or compositions are useful as
XX CC vaccines for treating or preventing Chlamydial infections including
XX CC pelvic inflammatory disease (which results in tubal obstruction and
XX CC infertility in women), male infertility, ocular infection (which may
XX CC cause blindness), acute respiratory tract infections, atherosclerosis, or
XX CC coronary heart disease. The present sequence is used in the
XX CC exemplification of the present invention.
XX XX
XX Sequence 660 AA;
XX SQ

```

Query Match      28.1%  Score 894;  DB 7;  Length 660;
Best Local Similarity 34.6%;  Pred. No. 2.4e-48;
Matches 233;  Conservative 126;  Mismatches 262;  Indels 52;  Gaps 17

QY      6  GPGCIDETRTPPADLSAQGLSAANKSAEORIAAEAKPKESKTDVSVERMSILRSAY 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      11  GPGVSSNCSNMPIINGQIASNETKSTYKTA-----SEASP--SASSSVSSWFLSSAK 63

```



```

QY      66  NALMSIADKLGITASSNSSSSSTGRS--ADVISTTATATPPPPPTDDYKTOAQTYAIDFTS 124
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      64  NALISURD--AALNKNSPTDLSQLEASTSTSTVTRVAAKQYDEAKSNFDTAKSGLENA 121
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      125  TSLADITQAAALVSIQDAVTNIKFAATD-----EET-----AIAAEWETKNADAVKVGQ 173
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      122  KTLAEVETKMDLMAALQDWERLANSDPSNNHTEEVNNIKKALEAKQDITD-----K 173
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      174  ITELAKYASDNQAILDSGLKLSFDLLQAALQSVANNKKAEBLLKEMQDNVPVPGKTPA 233
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      174  LMKLVTLQNKSTLEVLKTTDPSADQIPAINSOLEINKRSADQIIKDLFRQNT---SYEA 230
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      234  IAQSLVDQDQATATQIEKDGNAIRDYAFAGQNASGAVENAKSNNSISNTDSAKAAIATAK 293
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      231  VLTNAGEVIKASSEAGIKLQALQSLVDAGDQSAQVLAQQAQNSPDNTAAKELIDAE 290
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      294  TQIAEAAQK---FPDSPILQEBQWIOAEKULKNIKPADGSDVPNGPTTVGSGKQOOGSS 350
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      291  TKVNELKQEBHTGLTDSPLVKKAEQISQAQKIQEIKPS--GSDIPTVGPS--GSAAGAGSA 348
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      351  IGSI-----RVSMLLDDAENETASILMSGFROMTHMENTNPDSOAAQOQLAQAARAA 403
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      349  AGALKSSNNSGRISILLDDVDNEMAALQGFPSMIEQFVNNVPATAKLOAMEAQLTAM 408
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      404  K---AAGDSDSAAALADAQKALEBAALGKAGQOQGIINALQTIASAAVVSAGVPPAAASSI 460
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      409  SDQLVGAQDELPAETQAIKDALAQAL--KOPSDAGLGATAMQGVAFAAKAVGGGSAGTAGTV 467
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      461  GSSVKQLYKT--SKTSGDYKTOISAGYDAYKSINDAYGRANDAFRDVNNVSTPALTR 518
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      468  QMNVKQLYKTAFTSSTSSSAAALSDGYGAYKTLNSLYSESS--GVQSAISQTANPALSR 526
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      519  SVPRARTARG--PEKTDQALARVIGNSRSTLGDVYSQVSGALOSVMQIIOSNPOANNEETR 577
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      527  SVSRSGIESQGRSADASQRAAEITIVRDSQTLGDVYSRLQVLDSLSTIVSNPQANOEEM 586
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      578  QKLTSAVTKPPQFGYPYVQLSNDSTQKFTAKLESIPAEGRSRTAAETKALSFEETNSLFIQ 637
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      587  QKLTASISKAPFGYPVQNSADSLQKFAAQULEREFDGERSLARSQENAFKQPAFIIQ 646
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      638  VLVNIGSLYSGLV 650
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      647  VLVNIASLFSGLV 659
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 15

AAE13695	
ID	AAE13695 standard; protein; 361 AA.
XX	
XX	
AC	AAE13695;
XX	
DT	02-FEB-2001 (first entry)
XX	
XX	
DE	Chlamydia sp. protein # 6.
XX	
KW	Chlamydial infection; sexually transmitted disease;
KW	pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW	trachoma; blindness; acute respiratory tract infection; atherosclerosis;
XX	coronary heart disease; antibacterial.
XX	
OS	Chlamydia sp.
XX	
PN	WO200034483-A2.
XX	
PD	
XX	15-JUN-2000.
XX	
PF	08-DEC-1999; 99WO-US029012.
XX	
XX	
PR	08-DEC-1998; 98US-00208277.
PR	08-APR-1999; 99US-00288594.
PR	01-OCT-1999; 99US-00410568.
PR	22-OCT-1999; 99US-00426571.

XX (CORI-) CORIXA CORP.
XX
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
XX WPI; 2000-431303/37.
XX
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX PT comprises immunogenic portion of Chlamydia antigen, which comprises amino
XX PT acid sequence encoded by polynucleotide sequence.
XX
XX
XX Disclosure; Page 251-253; 256pp; English.
XX
XX
XX The present invention relates to new nucleic acid sequences and the
XX CC proteins encoded by the nucleic acid sequences. The encoded proteins
XX CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX CC proteins are useful for the serodiagnosis and treatment of Chlamydia
XX CC infection. Chlamidia are intracellular bacterial pathogens that are
XX CC responsible for a wide variety of human infections. C. trachomatis
XX CC infection is one of the most common sexually transmitted diseases and
XX CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX CC and infertility. Trachoma due to ocular infection with C. trachomatis is
XX CC the leading cause of preventable blindness worldwide. C. pneumonia is a
XX CC major cause of acute respiratory tract infections in humans and is also
XX CC thought to play a role in the pathogenesis of atherosclerosis and
XX CC coronary heart disease. The present sequence is a protein isolated in the
XX CC present invention
XX CC
XX Sequence 361 AA;
XX SQ

[illegible]

Search completed: March 23, 2004, 19:27:37
Job time : 62.3737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:53:56 ; Search time 17.7616 Seconds

(without alignments)

3525.629 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MVNPIGPGFIDETERTPPAD.....SLFIOQLVNIIGSLGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	3187	100.0	651	2	D72042	conserved hypothetical
2	3187	100.0	651	2	E86581	CHLPN 76 kDa homol
3	2205	69.2	715	2	I40729	hypothetical 76K p
4	899.5	28.2	647	2	G71490	hypothetical prote
5	232.5	7.3	2055	2	T31110	extracellular matr
6	232	7.3	1822	2	S33441	EF protein - Strep
7	223.5	7.0	2481	2	D90011	FmtB protein [impo
8	221.5	7.0	971	2	B90835	probable tail fibe
9	221.5	7.0	973	2	C85693	probable membrane
10	220	6.9	6713	2	B93921	hypothetical prote
11	210	6.6	1122	2	G64887	probable tail fibe
12	210	6.6	2271	2	F90073	hypothetical prote
13	208.5	6.5	4776	2	E95206	cell wall surface
14	199	6.2	1365	2	T30822	lmp1 protein - Mv
15	196.5	6.2	1829	2	T24583	hypothetical prote
16	196.5	6.2	2232	2	T34434	hypothetical prote
17	192	6.0	1156	2	T34852	probable secreted
18	190.5	6.0	1063	2	D86731	hypothetical prote
19	189	5.9	2155	2	AD2742	conserved hypothet
20	189	5.9	2155	2	C97523	hypothetical prote
21	187	5.9	1147	2	T35781	hypothetical prote
22	186.5	5.9	1306	2	S23370	MS22 protein - yea
23	183	5.7	2186	2	H89960	hypothetical prote
24	182.5	5.7	641	2	C92206	probable acceptin c
25	182.5	5.7	1238	2	T03465	methyl-accepting c
26	182.5	5.7	2541	2	S11661	tal in - mouse
27	182	5.7	1269	2	F84730	probable myosin he
28	178.5	5.6	1566	2	A43607	cell surface antig
29	177	5.6	2090	2	S26058	probable transform

ALIGNMENTS

RESULT 1

D72042
conserved hypothetical protein CP0018 [imported] - Chlamydomophila pneumoniae (strains CM N/Altermate names: chlpn 76 kda homolog 1 (ct622); hypothetical protein CPn0728
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: D72042; D81623
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammell, C.; Fan, J.; Olinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72042
A:Molecule type: DNA
A:Residues: 1-651 <ARN>
A:CROSS-references: GB:AB001654; GB:AB001363; NID:g4377031; PIDN:AAI18867.1; PID:g43770
A:Experimental source: Strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10694935
A:Accession: D81623
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-651 <REA>
A:CROSS-references: GB:AB002165; GB:AB002161; NID:g7188948; PIDN:AAF37914.1; PID:g71889
A:Experimental source: strain AR39, HL cells
C:Comment: This sequence was originally identified as homologous to part of a sequence PIR:H71490).
C:Genetics:
A:Gene: CPn0728; CP0018

Query Match 100.0%; Score 3187; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 1.9e-148;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MVNPIGPGFIDETERTPPADLSAAGLEASANKSAEQAQRIAGAAKPKESKTSVERWSI	60
DB	1	MVNPIGPGFIDETERTPPADLSAAGLEASANKSAEQAQRIAGAAKPKESKTSVERWSI	60
QY	61	LRSAVNALMSLADKLGIASSNSSSTSRSDVDSTTATPTPTPTTDFDYKTQQTAYDT	120
DB	61	LRSAVNALMSLADKLGIASSNSSSTSRSDVDSTTATPTPTPTTDFDYKTQQTAYDT	120
QY	121	IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAAAEWETKNADAVKVAQITELAKY	180
DB	121	IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAAAEWETKNADAVKVAQITELAKY	180
QY	181	ASDNCAILLDSGLKLTSDLLQALLOQSVANNKKAELLKEMODNPVVPKTPAAQSLVD	240
DB	181	ASDNCAILLDSGLKLTSDLLQALLOQSVANNKKAELLKEMODNPVVPKTPAAQSLVD	240

QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
 DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
 QY 301 KKFPPSPILQAEQWVIAEQKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSIRVSMML 360
 DB 301 KKFPPSPILQAEQWVIAEQKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAAKAAGDSDSAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAAKAAGDSDSAAALADAQK 420
 QY 421 ALBAALGKAGQOQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 DB 421 ALBAALGKAGQOQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 QY 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 DB 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 QY 541 ISGNSRTLDGVYSQVSALQSMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPVQLSND 600
 DB 541 ISGNSRTLDGVYSQVSALQSMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPVQLSND 600
 QY 601 STQKFIKLESFPAEGSRTAAEIKALSPETNSLFIQQVLVNIIGSLYSYLQ 651
 DB 601 STQKFIKLESFPAEGSRTAAEIKALSPETNSLFIQQVLVNIIGSLYSYLQ 651

RESULT 2

CHLPN 76 kDa homolog 1 (CT622) [imported] - Chlamydomophila pneumoniae (strain J138)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: E86581
 R:Shirai, M.; Hirakawa, H.; Kimoko, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: E86581
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-651 <STO>
 A:Cross-references: GB:BA000008; NID:G6979100; PIDM:BA098935.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: Cpj0728

Query Match 100.0%; Score 3187; DB 2; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.9e-148;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
 DB 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPTPTTDDYKTAQATYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPTPTTDDYKTAQATYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTATDETAIAAEWETKNADAVKVGQAQITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTNIKDTATDETAIAAEWETKNADAVKVGQAQITELAKY 180
 QY 181 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 DB 181 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
 DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
 QY 301 KKFPPSPILQAEQWVIAEQKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSIRVSMML 360

DB 301 KKFPPSPILQAEQWVIAEQKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAAKAAGDSDSAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAAKAAGDSDSAAALADAQK 420
 QY 421 ALBAALGKAGQOQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 DB 421 ALBAALGKAGQOQILNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
 QY 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 DB 481 QISAGYDAYKSIINDAYGRANDATRDVINNVSTPALTRSPRARTARGPEKTDQALARV 540
 QY 541 ISGNSRTLDGVYSQVSALQSMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPVQLSND 600
 DB 541 ISGNSRTLDGVYSQVSALQSMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPVQLSND 600
 QY 601 STQKFIKLESFPAEGSRTAAEIKALSPETNSLFIQQVLVNIIGSLYSYLQ 651
 DB 601 STQKFIKLESFPAEGSRTAAEIKALSPETNSLFIQQVLVNIIGSLYSYLQ 651

RESULT 3

140729
 hypoethical 76K protein - Chlamydomophila pneumoniae (strain AR39)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
 C:Accession: I40729
 R:Perez-Melgosa, M.; Kuo, C.
 Infect. Immun. 62, 880-886, 1994
 A:Title: Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-kDa
 A:Reference number: I40729; MUID:94156481; PMID:7509320
 A:Accession: I40729
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-715 <RES>
 A:Cross-references: GB:L23921; NID:G435961; PIDN:AAA23117.1; PID:G435962
 A:Experimental source: strain AR-39
 C:Comment: This is the hypothetical translation of a sequence that was reported as two

Query Match 69.2%; Score 2205; DB 2; Length 715;
 Best Local Similarity 99.6%; Pred. No. 1.9e-100;
 Matches 452; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 60
 DB 257 LVNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTDVSVERWSI 316
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPTPTTDDYKTAQATYDT 120
 DB 317 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPTPTTDDYKTAQATYDT 376
 QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTATDETAIAAEWETKNADAVKVGQAQITELAKY 180
 DB 377 IFTSTSLADIQAALVSLQDAVTNIKDTATDETAIAAEWETKNADAVKVGQAQITELAKY 436
 QY 181 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 DB 437 ASDNQAILDSLGLKTSFDLLQALLOSANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 496
 QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
 DB 497 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 556
 QY 301 KKFPPSPILQAEQWVIAEQKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSIRVSMML 360
 DB 557 KKFPPSPILQAEQWVIAEQKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSIRVSMML 616
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAAKAAGDSDSAAALADAQK 420
 DB 617 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAAKAAGDSDSAAALADAQK 676

RESULT 6

EF protein - Streptococcus suis
C:Species: Streptococcus suis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S33441
R:Smith, H.E.; Reek, F.H.; Vecht, U.; Gielkens, A.L.J.; Smits, M.A.
Submitted to the EMBL Data Library, May 1993
A:Description: Repeats in an extracellular protein of weak-pathogenic strains are absent
A:Reference number: S33441
A:Accession: S33441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1822 <SMI>
A:Cross-references: EMBL:X71880; NID:g298031; PIDN:CAA50714.1; PID:g298032

Query Match 7.3%; Score 232; DB 2; Length 1822;
Best Local Similarity 21.4%; Pred. No. 0.0014;
Matches 166; Conservative 127; Mismatches 289; Indels 192; Gaps 34;
QY 10 IDERTPTPADLSAOGLEASAAKSAEACRIAGAEAKPKESKTDV---ERWSILRSVAVN 66
DB 1072 IDDNENLPDE-----KESAKNAVEEAHVATA-AIDKASTPDVAVVEEDKGV--AAIN 1122
QY 67 ALMSLADKLGIASSNSSSTSRSD-----VDSSTTAT-----APTPTPTPDYKT 112
DB 1123 LITAKADAGVIAAKLADEIKLEBKQAEAKAIDASTWNEEKAIAKALQDVVDKGA 1182
QY 113 Q-----AQATYDITFTSLADIQAALV-----SLQAVYNIKDTAATDEETAIAAE-WETK 163
DB 1183 ELEDAAATVATNIHEATTTEKAKAEALAGEKSLTDTGKEARDAVBLAKDKELAKEAIRTE 1242
QY 164 NADAVKVGQAQITELAKYA-----SDNQA-----ILDSLGKLTSPDL 199
DB 1243 EEEATKIVEKAEADTRKALDENPNLSDEKQAEIKKLTDAVAKTLATIRNADKKTQ--- 1299
QY 200 LQAALLQSVANNKAAEALLKEMQDNPVPGKTPAIQAQSLVDQDTA-TQIEKDGNAIRD 258
DB 1300 -EAERQAALADLEKAKETQK-IADKAAIDRLTILVKDGELEATKQDAKNKIAKOAAAKAE 1357
QY 259 AYFAQONASGA-----VENAKSNNSIS-----NIDSAKAAIA 290
DB 1388 AIASPNLITDAEKKTTTDAVDAEAVAKANDASATSPADVQKBEADGVAIAAEVDLDAK 1417
QY 291 -TAKTQIAE---AQKPPDSPILOAEQMVTOAEKDLKNTKPADGSDVPNPGTTVGGSK 345
DB 1418 QDAKNKIADAAAKAEAGISNPNLDAEKKITFDVDAEAVAKANDASATSPADV--QK 1475
QY 346 QGSSSIGSVSMV---LDDAENETASILMSGFRQMIHMFNTENPDSQAQOEALAAQARA 402
DB 1476 EBDAGVAAIAEDVLDAAKQDAKNKIATKESDAAKSAIDANPNLITDAEKESAKKAVDADAKA 1535
QY 403 AKAGDDSAALALADAOKALEALGKAGQ-----QOQILNALGQIASAA 446
DB 1536 ATDAID--ASTSPVEAQAEDKGVGSIADQVLDAAKQDAKNKIATKESAKKAEADANPNL 1593
QY 447 -----VVSAGVPPAAASS-----IGSSVKQLYKTKSGTSDYKT 480
DB 1594 SDAEKEASKKAVDADAKATTDAIDASTSPVEAQAEDKGVGSIADQVLDAAK---QDAKN 1650
QY 481 QISAGYDAYKSINDAYGR---ARNDAITRVIN---NVSTPALTRSVPRPAREGEFE-KT 533
DB 1651 KIAKESDAKSAIDANPNLITDAEKESAKKAVDADAKAATDAIDAST--SPVEAQAEDK 1708
QY 534 DQALARVIGNSRITLGDVYVSQVS-ALQSVMOIIOQSNP-----QANNEEIRQKLTSAVT 585
DB 1709 VGAIAKIDILDAKQ--DAKNKIATKESAKSVIDSNPNLITDAEKAASKSEIDKAVEAIV 1766
QY 586 -----KPPQFGYPPYQLSNDSTQKFIKLESLFAEGS 617
DB 1767 LINGVRTYQLEBKIKLPMALIKPAKVTTPVD--PNNITEKELIARIKAFLENN 1819

RESULT 7

D90011
FMTS protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D90011
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu-
ma, A.; Mizutani-Ti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D90011
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2481 <XUR>
A:Cross-references: GB:BA000018; PID:g13701961; PIDN:BA043253.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: fmcB(mrp)

Query Match 7.0%; Score 223.5; DB 2; Length 2481;
Best Local Similarity 22.1%; Pred. No. 0.0056;
Matches 163; Conservative 129; Mismatches 293; Indels 153; Gaps 31;
QY 4 PIGPGPIDE-----TERTPPADLSAOGLEASAAKNSA-EACRIAGAEAKPKESKTDSDS--- 54
DB 763 PLNPDTTNEEVAEAIERINAAKVS--GVKAIBATTTAODLERVKNNEEIKFIENTIDSTQT 820
QY 55 -----VERWSILRSVAVNALMSLADKLGIASSNSSSTSRSD-----VDSSTTATAPT 101
DB 821 KMDAYKEVROAATARKAQNATVSNATDEEVAEANAADAACTEGHLHIOVVKSQEVADT 880
QY 102 PP-----PPFPDDYKTOAQATYDT-----IFTSLSADIQAALVSL--- 137
DB 881 KAKVLDKINAIQTOAKVKPAAD---TEVENAYNTRKQEIQNSNASTTTEEKAAVTELDK 937
QY 138 -QDAVTNIKDTAAYD-----EETAIAAEWETKNADAVKGAQITELAKYASDNQAILDS 190
DB 938 QEBARTNL-DAANTNSDVTTAKNGIAAINQVQAATTKKSDAK-AETQAQASERKTAIEA 995
QY 191 LGKLT-----SFDLQAAALLQSVANNKA-----AEILLKEMQDNPVPGKT--PAI 234
DB 996 MNDSTTEEQAAKDKVQAVVTANADIDNAFANTDVDNAKTNEATIAATITFDANVRPA 1055
QY 235 AQLVQDQDTATQIE-KQGNARIDAYFAGN-----ASGAVENAKSNNSI----- 279
DB 1056 KQAIADKVOAQETAIDANNNGSTTEEKAAKQOVQTEKTAADAIDAASHNVEVEBAKNAE 1115
QY 280 -----SNIDSAAKAIAT-----AKTOIAEAKKFPDPSPIQEA--EQWVIOAKD 322
DB 1116 IAKTEATQPTTTKDNKAQAIATKANERKTAIAGTQITAEIAAANADVNAVQAN-- 1173
QY 323 LKNITKPADG-SDVENPGTTVGGSKQSGSISIRVSMILLDDAENETASILMSGFRQMIH 381
DB 1174 -SNIEAANSQNDVDQAKTTGETSIDQVTPVNNKAT-----ARNEITAILNNKLQEIQT 1227
QY 382 FNTENPDSQAQOEALAAQAAKAAAGDDSA--ALADAQKALEALGKAGQOQG 434
DB 1228 PDATDEBKQAADAE--ANTENKANQAIASAATNAQVDEAKANAEEAINAVTPKVVKKQA 1285
QY 435 ILNALGQI-ASAAVWSAGVPPAAASSIGSSVKQLYKTKSGTSDYKTQISAGYDAYKSIN 493
DB 1286 AKDEIDQLQATQTVINNDQNAINEEKEAIIQOL---ATAVTDKNNITA----- 1332
QY 494 DAYGRANDATRVINNVSTPALTRSVPRPAREGEPEKTDQALARVISGNSRITLGDVYS 553
DB 1333 -----ATDDNGVDTPAKDAGKNSIQSTQTPATAVKNNAKNEVDQAVTTQNAIDNTTGATTE 1387
QY 554 QVSALQSVMOIIOQSNPQANNEERQKLTSAVTTPQFGYPPVQ--LSNDSTQKFIKLESL 612
DB 1388 EKNAAKOL--VLKAKAKAYQDILNQAQNTNDVTQIKDAQVADIQGITADTTIKDVAK---- 1441

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:24:24 ; Search time 39.1412 Seconds
(without alignments)
3262.843 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204
Sequence: 1 MWNIGPGIDETERTTPAD.....QGILNALQIASAAVWSAGV 452

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq_29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2204	100.0	452	3	AAy71956 3'-trunca
2	2204	100.0	490	3	AAy71957 C. pneumo
3	2204	100.0	651	3	AAy71954 Chlamydia
4	2204	100.0	651	4	AAU38899 C. pneumo
5	2204	100.0	651	5	ABb94277 Chlamydia
6	2204	100.0	651	5	ABb90532 Chlamydia
7	2204	100.0	651	7	ADd42665 Chlamydia
8	2182	99.0	478	2	AAy35358 Chlamydia
9	1866	84.7	583	3	AAy71955
10	461.5	20.9	647	7	ABb94283 Chlamydia
11	461.5	20.9	647	7	ADD43866 Chlamydia
12	453	20.6	660	7	AAu38921 C. tracho
13	453	20.6	660	7	ADd42726 Chlamydia
14	289.5	13.1	350	2	AAy37571 Chlamydia
15	231.5	10.5	361	3	AAb13695 Chlamydia
16	231.5	10.5	361	4	AAg83263 Chlamydia
17	231.5	10.5	361	5	ABb94234 Chlamydia
18	219	9.9	1463	5	AAE20110 Lactobaci
19	197	8.9	971	7	ADc00627 Enterohae
20	190	8.6	2478	4	AAU37374 Staphyloc
21	190	8.6	2478	4	AAU34320 Staphyloc
22	190	8.6	2478	6	ABj19002 Pathogen
23	190	8.6	2478	6	ABm71899 Staphyloc
24	189.5	8.6	1831	6	ABu43109 Protein e
25	189.5	8.6	1870	6	ABj19019 Pathogen

26	188	8.5	603	6	ABU44080 Protein e
27	186.5	8.5	790	6	ABU47218 Protein e
28	184	8.3	1122	6	ABU14859 Protein e
29	183.5	8.3	1822	2	AAr27745 Extracell
30	181	8.2	2481	7	ABU15838 Protein e
31	180	8.2	2481	7	ABR62804 Methicill
32	180	8.2	2137	5	ABP39618 Staphyloc
33	178	8.1	1178	6	ABj26182 Aspergill
34	176	8.0	1924	6	ABU43922 Protein e
35	173	7.8	2271	6	ABU16000 Protein e
36	172.5	7.8	812	6	ABU47318 Protein e
37	172	7.8	2016	6	AAE36891 Plectreur
38	172	7.8	2344	4	AAU37120 Staphyloc
39	171.5	7.8	1879	5	AAE20111 Lactobaci
40	170.5	7.7	2368	4	AAU34139 Staphyloc
41	170.5	7.7	2368	4	AAU36796 Staphyloc
42	169.5	7.7	2261	6	ABj18914 Pathogen
43	168.5	7.6	2398	6	ABU42252 Protein e
44	168.5	7.6	6281	4	AAU37403 Staphyloc
45	168.5	7.6	9535	6	ABM73008 Staphyloc

ALIGNMENTS

RESULT 1
AAy71956
ID AAY71956 standard; protein; 452 AA.
XX AC AAY71956;
XX DT 26-MAR-2001 (first entry)
XX DE 3'-truncated Chlamydia pneumoniae 76 kDa protein.
XX KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; truncation mutant; mutein.
XX OS Chlamydophila pneumoniae.
XX OS Synthetic.
XX EN WO200066739-A2.
XX PD 09-NOV-2000.
XX PF 03-MAY-2000; 2000WO-CA000511.
XX PR 03-MAY-1999; 99US-0132270P.
XX PR 30-JUN-1999; 99US-0141276P.
XX PA (AVET) AVENTIS PASTEUR LTD.
XX PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX DR WPI; 2000-687542/67.
XX DR N-PSDB; AAD02065.
XX PT Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
XX for vaccinating against Chlamydia infections.
XX PS Claim 16c; Page 104-106; 112pp; English.
XX CC The present sequence is 3'-truncated Chlamydia pneumoniae 76 kDa protein.
XX CC C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
XX CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
XX CC tract disease, bronchitis, sinusitis and acute respiratory disease such
XX CC as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
XX CC auscultation). C. pneumoniae sequence is also used as vaccines for
XX CC immunising humans against diseases caused by C. pneumoniae
XX SQ Sequence 452 AA;

applicants

Query Match 100.0%; Score 2204; DB 3; Length 452;
 Best Local Similarity 100.0%; Pred. No. 2.9e-146;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITGPIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVERWSI 60
 DB 1 MNPITGPIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVERWSI 60
 QY 61 LRSVAVNMLSLADKLGIIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
 DB 61 LRSVAVNMLSLADKLGIIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSLGLKTSFDLLQALQLQSVANNKAAELLKEMQDNVPVPGKTPAIQAQSLVD 240
 DB 181 ASDNQAILDSLGLKTSFDLLQALQLQSVANNKAAELLKEMQDNVPVPGKTPAIQAQSLVD 240
 QY 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPDPSPIQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPDPSPIQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFROMIHMTENTPDSQAQQLAQAARAAKAGDSDSAAALADAK 420
 DB 361 DDAENETASILMSGFROMIHMTENTPDSQAQQLAQAARAAKAGDSDSAAALADAK 420
 QY 421 ALEAALGKAGQGGILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQGGILNALGQIASAAVVSAGV 452

RESULT 2
 AAY71957
 ID AAY71957 standard; protein; 490 AA.
 XX
 AC AAY71957;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE C. pneumoniae 76 kDa protein truncation mutant fusion protein.
 XX
 KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; fusion protein; truncation mutant;
 KW mutin.
 XX
 OS Chlamydothila pneumoniae.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 453..490
 FT /note= "This part of the sequence is unrelated to the C.
 FT pneumoniae 76 kDa protein"
 XX
 PN WO200066739-A2.
 XX
 XX 09-NOV-2000.
 XX
 XX 03-MAY-2000; 2000WO-CA000511.
 XX
 XX 03-MAY-1999; 99US-0132270P.
 XX 30-JUN-1999; 99US-0141276P.
 XX
 XX (AVET) AVENTIS PASTEUR LTD.
 PA
 XX

Murdin AD, Comen RP, Wang J, Dunn P;
 WPI; 2000-687542/67.
 DR N-PSDB; AAD02066.
 XX
 Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
 PT for vaccinating against Chlamydia infections.
 XX
 Claim 33; Fig 3; 112pp; English.
 PS
 CC The present sequence is a fusion protein comprising a truncated Chlamydia
 CC pneumoniae 76 kDa protein and vector-encoded residues. C. pneumoniae 76
 CC kDa protein is used in the diagnosis, prevention and treatment of C.
 CC pneumoniae infections (e.g. pneumonia, upper respiratory tract disease,
 CC bronchitis, sinusitis and acute respiratory disease such as cough, sore
 CC throat, hoarseness, fever; and abnormal chest sounds on auscultation). C.
 CC pneumoniae sequence is also used as vaccines for immunising humans
 CC against diseases caused by C. pneumoniae
 XX
 SQ Sequence 490 AA;
 Query Match 100.0%; Score 2204; DB 3; Length 490;
 Best Local Similarity 100.0%; Pred. No. 3.2e-146;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPITGPIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVERWSI 60
 DB 1 MNPITGPIDETERTPPADLSAQGLEASANKSAEARIAGAKPKESKTDVERWSI 60
 QY 61 LRSVAVNMLSLADKLGIIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
 DB 61 LRSVAVNMLSLADKLGIIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSLGLKTSFDLLQALQLQSVANNKAAELLKEMQDNVPVPGKTPAIQAQSLVD 240
 DB 181 ASDNQAILDSLGLKTSFDLLQALQLQSVANNKAAELLKEMQDNVPVPGKTPAIQAQSLVD 240
 QY 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPDPSPIQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPDPSPIQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFROMIHMTENTPDSQAQQLAQAARAAKAGDSDSAAALADAK 420
 DB 361 DDAENETASILMSGFROMIHMTENTPDSQAQQLAQAARAAKAGDSDSAAALADAK 420
 QY 421 ALEAALGKAGQGGILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQGGILNALGQIASAAVVSAGV 452

RESULT 3
 AAY71954
 ID AAY71954 standard; protein; 651 AA.
 XX
 AC AAY71954;
 XX
 DT 12-SEP-2003 (revised)
 DT 26-MAR-2001 (first entry)
 XX
 XX Chlamydia pneumoniae 76 kDa full-length protein.
 DE
 KW 76 kDa protein; bactericidal; diagnosis; prevention; treatment;
 KW pneumonia; upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation.

XX OS Chlamydothila pneumoniae.
XX PN WO200066739-A2.
XX AC
XX XX
XX PD 09-NOV-2000.
XX PF 03-MAY-2000; 2000WO-CA000511.
XX PR 03-MAY-1999; 99US-0132270P.
XX PR 30-JUN-1999; 99US-0141276P.
XX PA (AVET) AVENTIS PASTEUR LTD.
XX PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX DR WPI; 2000-687542/67.
XX DR N-PSDB; RAD02063.
XX XX
XX PT Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
XX PT for vaccinating against Chlamydia infections.
XX PS Claim 16a; Fig 1; 112pp; English.
XX XX
XX CC The present sequence is Chlamydia pneumoniae full-length 76 kDa protein.
XX CC C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
XX CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
XX CC tract disease, bronchitis, sinusitis and acute respiratory disease such
XX CC as cough, sore throat, hoarseness, fever; and abnormal chest sounds on
XX CC auscultation). C. pneumoniae sequence is also used as vaccines for
XX CC immunising humans against diseases caused by C. pneumoniae. (Updated on
XX CC 12-SEP-2003 to standardise OS field)
XX XX
XX SQ Sequence 651 AA;
Query Match 100.0%; Score 2204; DB 3; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.6e-146; Indels 0; Gaps 0;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVNPIGPGIDETERTPPADLSAAGLEASAAKSAEQAQRIAGAEAKPKSKTDSVERWSI 60
Db 1 MVNPIGPGIDETERTPPADLSAAGLEASAAKSAEQAQRIAGAEAKPKSKTDSVERWSI 60
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
Db 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETAEAEWETKNADAVKGAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETAEAEWETKNADAVKGAQITELAKY 180
QY 181 ASDNQAILDLSLGLTSPDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
Db 181 ASDNQAILDLSLGLTSPDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
QY 241 QTDATATQIEKGNARIDAVFAGNAGAVENAKSNNSINISIDAKAIAIATAKTQIAEAQ 300
Db 241 QTDATATQIEKGNARIDAVFAGNAGAVENAKSNNSINISIDAKAIAIATAKTQIAEAQ 300
QY 301 KKFPPDSPILOAEQWVIOAEKDLKNIKPADGSDVNPCTTVGSKQGGSSIGSRVSMLL 360
Db 301 KKFPPDSPILOAEQWVIOAEKDLKNIKPADGSDVNPCTTVGSKQGGSSIGSRVSMLL 360
QY 361 DDAENETASILMSGFRMIHMFNTENPDSQAQOELAAQAPAAKAGDSDAAAALADAQK 420
Db 361 DDAENETASILMSGFRMIHMFNTENPDSQAQOELAAQAPAAKAGDSDAAAALADAQK 420
QY 421 ALEAALGKAGQOQOQILNALGQIAAAVVSAGV 452
Db 421 ALEAALGKAGQOQOQILNALGQIAAAVVSAGV 452
RESULT 4

AAU38899
ID AAU38899 standard; protein; 651 AA.
XX AC
XX AC AAU38899;
XX XX
XX DT 11-SEP-2003 (revised)
XX DT 16-JAN-2002 (first entry)
XX XX
XX DE C. pneumoniae CT622 homologue CPn0728.
XX XX
XX KW Chlamydia; sexually transmitted disease; PID; antibacterial;
XX KW pelvic inflammatory disease; antigen; trachoma; gynecological;
XX KW acute respiratory tract infection; atherosclerosis; male infertility;
XX KW coronary heart disease.
XX XX
XX OS Chlamydothila pneumoniae.
XX XX WO200181379-A2.
XX PN
XX PD 01-NOV-2001.
XX PF 23-APR-2001; 2001WO-US013081.
XX PR 21-APR-2000; 2000US-0198953P
XX PR 20-JUL-2000; 2000US-0219752P.
XX PA (CORI-) CORIXA CORP.
XX PI Bhatia A, Probst P, Stromberg EJ;
XX DR WPI; 2001-616771/71.
XX DR N-PSDB; AAS57031.
XX XX
XX PT New polynucleotide for treating Chlamydia infections encodes a
XX PT polynucleotides containing an immunogenic portion of a Chlamydia antigen.
XX PS Disclosure; Page 161-162; 208pp; English.
XX XX
XX CC The invention relates to isolated polynucleotide encoding at least a
XX CC partial Chlamydia protein which is an antigenic fragment, or the
XX CC complements, fragments, homologues and variants, and antibodies raised
XX CC against the antigenic proteins (or fragments). The nucleic acids,
XX CC proteins and antibodies are used to diagnose and treat Chlamydia
XX CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
XX CC disease (PID), acute respiratory tract infection, trachoma,
XX CC atherosclerosis and coronary heart disease) in a patient, and in the
XX CC treatment of male infertility. The compounds of the invention are also
XX CC useful for detecting the presence of Chlamydia in a patient, and
XX CC stimulating and/or expanding T cells specific for a Chlamydia protein.
XX CC The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-
XX CC 2003 to standardise OS field)
XX SQ Sequence 651 AA;
Query Match 100.0%; Score 2204; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.6e-146; Indels 0; Gaps 0;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVNPIGPGIDETERTPPADLSAAGLEASAAKSAEQAQRIAGAEAKPKSKTDSVERWSI 60
Db 1 MVNPIGPGIDETERTPPADLSAAGLEASAAKSAEQAQRIAGAEAKPKSKTDSVERWSI 60
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
Db 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTTDDYKTAQATYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETAEAEWETKNADAVKGAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNPKDTAETAEAEWETKNADAVKGAQITELAKY 180
QY 181 ASDNQAILDLSLGLTSPDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
Db 181 ASDNQAILDLSLGLTSPDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240

QY 241 QTATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPDSPILQEAQOMVIOAEKDLKNIKPAGSDVPNGTTVGSKQGGSSIGSIRVSMML 360
 Db 301 KKFPDSPILQEAQOMVIOAEKDLKNIKPAGSDVPNGTTVGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAGDDSAALADAQK 420
 Db 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAGDDSAALADAQK 420
 QY 421 ALAALGKAGQCGGILNALGQIASAAVVSAGV 452
 Db 421 ALAALGKAGQCGGILNALGQIASAAVVSAGV 452

RESULT 5

ABB94277 standard; protein; 651 AA.
 XX ABB94277;

AC 29-AUG-2003 (revised)
 DT 05-JUN-2002 (first entry)

Chlamydia pneumoniae protein sequence SEQ ID NO:405.

Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
 antibacterial; immunostimulant; immune response;
 Chlamydia-specific T-cell response.

Chlamydia pneumoniae.

WC200208267-A2.

31-JAN-2002.

20-JUL-2001; 2001WO-US023121.

20-JUL-2000; 2000US-00620412.

23-APR-2001; 2001US-00841132.

(CORI-) CORIYA CORP.

Fling SP, Skeiky YAW, Probst P, Bhatia A;

WPI; 2002-179901/23.

Novel compositions comprising Chlamydia Cap1 protein and its use in the
 treatment of Chlamydia infection.

Disclosure; Page 370-372; 537pp; English.

The present invention describes compositions comprising a Chlamydia Cap1
 protein and methods for the diagnosis and therapy of Chlamydia infection.
 Chlamydia DNA and protein sequences from the present invention can have
 antibacterial and immunostimulant activities, and can be used in
 vaccines. Compounds from the present invention can be used for eliciting
 an immune response, specifically stimulating a Chlamydia-specific T-cell
 response or inhibiting the development of a Chlamydia infection in an
 animal. Methods from the present invention can be used for detecting the
 presence of Chlamydia in a patient; to stimulate and/or expand T cells
 specific for a Chlamydia protein; and for treatment of a Chlamydia
 infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
 sequences used in the exemplification of the present invention. (Updated
 on 29-AUG-2003 to standardise OS field)

Sequence 651 AA;

Query Match

Best Local Similarity 100.0%; Score 2204; DB 5; Length 651;

Pred. No. 4.6e-146;

Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MNPPIGPGPIDETERPPADLSAQGLEASAANKSAEQAQIAGAAKPKSKTDSVERWSI 60
 Db 1 MNPPIGPGPIDETERPPADLSAQGLEASAANKSAEQAQIAGAAKPKSKTDSVERWSI 60
 QY 61 LRSVNNALMSLADKLGIIASSNSSSTSRSDVSDTTATATPTPPPTFDDYKTAQATAYDT 120
 Db 61 LRSVNNALMSLADKLGIIASSNSSSTSRSDVSDTTATATPTPPPTFDDYKTAQATAYDT 120
 QY 121 IFTSLSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 Db 121 IFTSLSLADIQAALVSLQDAVNIKDTATDEETAIAAEWETKNADAVKVGAIITELAKY 180
 QY 181 ASDNQAILDSLGLKITSFDLLQAALLQSVANNKKAELLKEMODNPVVPKTPAIAGSLVD 240
 Db 181 ASDNQAILDSLGLKITSFDLLQAALLQSVANNKKAELLKEMODNPVVPKTPAIAGSLVD 240
 QY 241 QTATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTATATQIEKGNIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPDSPILQEAQOMVIOAEKDLKNIKPAGSDVPNGTTVGSKQGGSSIGSIRVSMML 360
 Db 301 KKFPDSPILQEAQOMVIOAEKDLKNIKPAGSDVPNGTTVGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAGDDSAALADAQK 420
 Db 361 DDAENETASILMSGFQMIHMENTENPDSQAQOELAAQARAARAKAGDDSAALADAQK 420
 QY 421 ALAALGKAGQCGGILNALGQIASAAVVSAGV 452
 Db 421 ALAALGKAGQCGGILNALGQIASAAVVSAGV 452

RESULT 6

ABB90532 standard; protein; 651 AA.
 XX ABB90532;

AC 29-AUG-2003 (revised)

DT 29-JUL-2002 (first entry)

Chlamydia pneumoniae cp7033 protein, SEQ ID NO:13.

Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
 human respiratory disease; cardiovascular disease; atherosclerosis;
 coronary artery disease; carotid artery stenosis; myocardial infarction;
 cerebrovascular disease; aortic aneurysm; claudication; stroke;
 strain CWL029.

Chlamydia pneumoniae.

WO200202606-A2.

10-JAN-2002.

03-JUL-2001; 2001WO-IB001445.

03-JUL-2000; 2000GB-00016369.

11-JUL-2000; 2000GB-00017047.

21-JUL-2000; 2000GB-00017983.

07-AUG-2000; 2000GB-00019368.

18-AUG-2000; 2000GB-00020440.

14-SEP-2000; 2000GB-00022583.

10-NOV-2000; 2000GB-00027549.

22-DEC-2000; 2000GB-00031705.

(CHIR-) CHIRON SPA.

Ratti G, Grandi G;

DR WPI; 2002-154726/20.
DR N-PSDB; ABL91190.
XX
PT Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
PT preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX
PS Claim 1; Page 47; 364pp; English.
XX
CC Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CWL029) and ABL91184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
CC DNA probe assay or blotting techniques for determining Chlamydia
CC pneumoniae gene expression. The present sequence represents a
CC specifically claimed Chlamydia pneumoniae protein of the invention.
CC (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 651 AA;

Query Match 100.0%; Score 2204; DB 5; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.6e-146; Indels 0; Gaps 0;
Matches 452; Conservative 0; Mismatches 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60
Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60

QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTQATAYDT 120
Db 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTQATAYDT 120

QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180

QY 181 ASDNQAILDSLGLKLTASFLLQALQLQSVANNKAAELKEMQDNFVPGKTPAIAQSLVD 240
Db 181 ASDNQAILDSLGLKLTASFLLQALQLQSVANNKAAELKEMQDNFVPGKTPAIAQSLVD 240

QY 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKNNSISNTDSAKAIAATAKTOIAEAQ 300
Db 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKNNSISNTDSAKAIAATAKTOIAEAQ 300

QY 301 KKFPPSPILQEAQEMVIOAEKDLKNIKADGSDVNPPTTGVGSKQGGSSIGSIRVSMLL 360
Db 301 KKFPPSPILQEAQEMVIOAEKDLKNIKADGSDVNPPTTGVGSKQGGSSIGSIRVSMLL 360

QY 361 DDAENETASILMSGFRQTHMENTENPDSPQAQQLAQARAAKAGDDSAALADAAQK 420
Db 361 DDAENETASILMSGFRQTHMENTENPDSPQAQQLAQARAAKAGDDSAALADAAQK 420

QY 421 ALEAALGRAGQOQGIILNALGQIASAAVVSAGV 452
Db 421 ALEAALGRAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 7
ADD42665
ID ADD42665 standard; protein; 651 AA.
XX
AC ADD42665;
XX
DT 15-JAN-2004 (first entry)

XX Chlamydia pneumoniae antigen protein SEQ ID NO:78.
DE
XX
KW Chlamydial infection; Chlamydia; antibiotic; antiinflammatory;
KW antiinfertility; cardiant; antiarteriosclerotic; ophthalmological;
KW vaccine; gene therapy; immune response; pelvic inflammatory disease;
KW tubal obstruction; infertility; male infertility; ocular infection;
KW blindness; acute respiratory tract infection; atherosclerosis;
KW coronary heart disease.
XX
OS Chlamydia pneumoniae.
XX
XX WO2003041560-A2.
XX 22-MAY-2003.
XX 05-NOV-2002; 2002WO-US035624.
XX 06-NOV-2001; 2001US-00012246.
XX 05-DEC-2001; 2001US-00007493.
XX 15-JUL-2002; 2002US-00197220.
XX (CORI-) CORIXA CORP
XX
PI Bhatia A, Guérian J, Skeiky YAW, Maisonneuve JL, Barth B;
PI Probst P;
XX
DR WPI; 2003-441771/41.
DR N-PSDB; ADD42825.
XX
PT New DNA and proteins comprising a portion of a Chlamydia antigen, useful
PT for diagnosing or treating Chlamydial infections, particularly as
PT vaccines for treating or preventing Chlamydial infections, e.g. pelvic
PT inflammatory disease.
XX
PS Disclosure; SEQ ID NO 78; 275pp; English.
XX
CC The present invention describes compounds and methods for diagnosing and
CC treating Chlamydial infection. Chlamydia polynucleotide and protein
CC sequences have antibiotic, antiinflammatory, antiinfertility, cardiant,
CC antiarteriosclerotic and ophthalmological activities, and can be used in
CC vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,
CC compositions or methods from the present invention can be used for the
CC serodiagnosis or treatment of Chlamydial infections, particularly in
CC humans. The polynucleotides, proteins or compositions are particularly
CC useful for stimulating an immune response in a patient, or for
CC stimulating and/or expanding T cells specific for a Chlamydia protein.
CC Specifically, the polynucleotides, proteins or compositions are useful as
CC vaccines for treating or preventing Chlamydial infections including
CC pelvic inflammatory disease (which results in tubal obstruction and
CC infertility in women), male infertility, ocular infection (which may
CC cause blindness), acute respiratory tract infections, atherosclerosis, or
CC coronary heart disease. The present sequence is used in the
CC exemplification of the present invention.
XX
SQ Sequence 651 AA;

Query Match 100.0%; Score 2204; DB 7; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.6e-146;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60
Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEAOIAGAEAKPKESKTSVERWSI 60

QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTQATAYDT 120
Db 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDDYKTQATAYDT 120

QY 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVNTIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180

QY 181 ASDNQALDLSGLKLTSDLLQALLOSVANNKKAELLKEMQNDPVVPGKTPAIAQSLVD 240
 DB 181 ASDNQALDLSGLKLTSDLLQALLOSVANNKKAELLKEMQNDPVVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQ 300
 DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQ 300
 QY 301 KKPDPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOGSSIGSIRVSMML 360
 DB 301 KKPDPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOGSSIGSIRVSMML 360
 QY 361 DDENETASILMSGFRQMIHMENTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 DB 361 DDENETASILMSGFRQMIHMENTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 QY 421 ALEAALGKAGQQQGGILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQQQGGILNALGQIASAAVVSAGV 452

RESULT 8

AAV35358
 ID AAY35358 standard; protein; 478 AA.

XX AC AAY35358;

DT 17-OCT-2003 (revised)

DT 13-SEP-1999 (first entry)

XX Chlamydia pneumoniae surface exposed polypeptide.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 KW neutralising epitope.

OS Chlamydia pneumoniae.

XX PN W0927105-A2.

XX PD 03-JUN-1999.

XX PF 30-NOV-1999, 98WO-IB001890.

XX PR 21-NOV-1997; 97ER-00014573.

XX PR 04-NOV-1998; 98US-0107078P.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX DR WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae.

XX Page 1156; Disclosure; 1912pp; English.

CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)

XX SQ Sequence 478 AA;

Query Match

Best Local Similarity 99.0%; Score 2182; DB 2; Length 478;

DR N-PSDB; AAD02064.

Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAAANKSAEAQRIAGAEAKPKESKTSVERWSI 60
 DB 3 MWNPIGPGPIDETERTPPADLSAQGLEASAAANKSAEAQRIAGAEAKPKESKTSVERWSI 62
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATATPTPPPTFFDDYKTKQAQATYDT 120
 DB 63 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATATPTPPPTFFDDYKTKQAQATYDT 122
 QY 121 IFTSTSLADIQAALVSLQDAVNIKDTAATDEBETAAAEWETKNADAVKVGQITELAKY 180
 DB 123 IFTSTSLADIQAALVSLQDAVNIKDTAATDEBETAAAEWETKNADAVKVGQITELAKY 182
 QY 181 ASDNQAILDLSGLKLTSDLLQALLOSVANNKKAELLKEMQNDPVVPGKTPAIAQSLVD 240
 DB 183 ASDNQAILDLSGLKLTSDLLQALLOSVANNKKAELLKEMQNDPVVPGKTPAIAQSLVD 242
 QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQ 300
 DB 243 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATQIAEAQ 302
 QY 301 KKPDPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOGSSIGSIRVSMML 360
 DB 303 KKPDPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPFGTTVGGSKQOGSSIGSIRVSMML 362
 QY 361 DDENETASILMSGFRQMIHMENTENPDSQAQOELAAQARAAGDDSAALADAQK 420
 DB 363 DDENETASILMSGFRQMIHMENTENPDSQAQOELAAQARAAGDDSAALADAQK 422
 QY 421 ALEAALGKAGQQQGGILNALGQIASAAV 447
 DB 423 ALEAALGKAGQQQGGILNALGQIASAAV 449

RESULT 9
 AAY71955
 ID AAY71955 standard; protein; 583 AA.
 XX AC AAY71955;
 XX DT 26-MAR-2001 (first entry)
 XX DE 5'-truncated Chlamydia pneumoniae 76 kDa protein.
 XX KW 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
 KW upper respiratory tract disease; bronchitis; sinusitis;
 KW acute respiratory disease; cough; sore throat; hoarseness; fever;
 KW vaccine; immunisation; treatment; truncation mutant; mutein.
 XX OS Chlamydia pneumoniae.
 OS Synthetic.
 XX Key Location/Qualifiers
 XX FT Misc-difference 497
 XX FT Note= "Encoded by ACT"
 XX PN W0200066739-A2.
 XX PD 09-NOV-2000.
 XX PF 03-MAY-2000; 2000WO-CA000511.
 XX PR 03-MAY-1999; 99US-0132270P.
 XX PR 30-JUN-1999; 99US-0141276P.
 XX PA (AVET) AVENTIS PASTEUR LTD.
 XX PI Murdin AD, Omen RP, Wang J, Dunn P;
 XX DR WPI; 2000-687542/67.
 XX DR N-PSDB; AAD02064.

PT Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
PT for vaccinating against Chlamydia infections.

PS Claim 16b; Page 100-102; 112pp; English.

XX The present sequence is 5'-truncated Chlamydia pneumoniae 76 kDa protein.
CC C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
CC treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
CC tract disease, bronchitis, sinusitis and acute respiratory disease such
CC as cough, sore throat, hoarseness, fever, and abnormal chest sounds on
CC auscultation). C. pneumoniae sequence is also used as vaccines for
CC immunising humans against diseases caused by C. pneumoniae

XX Sequence 583 AA;

Query Match 84.7%; Score 1866; DB 3; Length 583;

Best Local Similarity 100.0%; Pred. No. 2.1e-122;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 MSLADKLGASNSSSSTRSADVDSTTATPTPPPTDDYKTAQAYDTIFTSTSLA 128

Db 1 MSLADKLGASNSSSSTRSADVDSTTATPTPPPTDDYKTAQAYDTIFTSTSLA 60

QY 123 DIQAALVSLQDAVNIKDTAATDEETAABWETKNADAVKVGQITELAKYASDQAIL 188

Db 61 DIQAALVSLQDAVNIKDTAATDEETAABWETKNADAVKVGQITELAKYASDQAIL 120

QY 189 DSLGKLTDFDLLOALLOSVAANNKAELLKEMQNDPVPVPGKTPAIAQSLVDQTDATQ 248

Db 121 DSLGKLTDFDLLOALLOSVAANNKAELLKEMQNDPVPVPGKTPAIAQSLVDQTDATQ 180

QY 249 LEKQGNALRDAYFAGQNASGAVENAKSNNSISNDSAKAATATAKTQTAEAQKPPDSP 308

Db 181 LEKQGNALRDAYFAGQNASGAVENAKSNNSISNDSAKAATATAKTQTAEAQKPPDSP 240

QY 309 LQEAQEMVIAEQKLNKIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 368

Db 241 LQEAQEMVIAEQKLNKIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 300

QY 369 SILMSGFRQMTLHMTENTPDSQAQQOELAAQARAAGDSDSAAALADAQKALEALGK 428

Db 301 SILMSGFRQMTLHMTENTPDSQAQQOELAAQARAAGDSDSAAALADAQKALEALGK 360

QY 429 AGQQGGIILNALGQIASAAVVSAGV 452

Db 361 AGQQGGIILNALGQIASAAVVSAGV 384

RESULT 10

ABB94283

ID ABB94283 standard; protein; 647 AA.

XX AC ABB94283;

XX 05-JUN-2002 (first entry)

XX Chlamydia trachomatis protein sequence SEQ ID NO:436.

XX Chlamydia trachomatis.

XX WO200208267-A2.

XX 31-JAN-2002.

XX 20-JUL-2001; 2001WO-US023121.

XX 20-JUL-2000; 2000US-00620412.

XX 23-APR-2001; 2001US-00841132.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

XX Chlamydia trachomatis.

PA (CORI-) CORIXA CORP.

XX Fling SP, Skeiky YAW, Probst P, Bhatia A;

XX MPI; 2002-179901/23.

XX Novel compositions comprising Chlamydia Cap1 protein and its use in the

XX treatment of Chlamydia infection.

XX Claim 34; Page 391-393; 537pp; English.

XX The present invention describes compositions comprising a Chlamydia Cap1

XX protein and methods for the diagnosis and therapy of Chlamydia infection.

XX Chlamydia DNA and protein sequences from the present invention can have

XX antibacterial and immunostimulant activities, and can be used in

XX vaccines. Compounds from the present invention can be used for eliciting

XX an immune response, specifically stimulating a Chlamydia-specific T-cell

XX response or inhibiting the development of a Chlamydia infection in an

XX animal. Methods from the present invention can be used for detecting the

XX presence of Chlamydia in a patient; to stimulate and/or expand T cells

XX specific for a Chlamydia protein; and for treatment of a Chlamydia

XX infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent

XX sequences used in the exemplification of the present invention

XX Sequence 647 AA;

QY Query Match 20.9%; Score 461.5; DB 5; Length 647;

Db Best Local Similarity 29.7%; Pred. No. 7.9e-24;

Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;

QY 2 VNPFGPIDETERTPPADLSAQGLEASAANKSAFAQRIAGAEAKPKESKTDSTVERWSIL 61

Db 15 MNPFGPIDETERTPPADLSAQGLEASAANKSAFAQRIAGAEAKPKESKTDSTVERWSIL 52

QY 62 RSANVALMSLADKLGASNSSSSTRS-ADVDTTATATPTPPPTDDYKTAQAYDT 120

Db 53 SSAGHALISLRD--AIIINKSSPTDSLQLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110

QY 121 IFTSTSLADTQAALVSLQDAVNIKDTAATDEE-TAJAAEWETKNADAVKVGQITELAK 179

Db 111 LENATTAEVETKMDLMAALQDMERLAKQAEVTRIKEALQEKQ---EVIDKLNQLVK 166

QY 180 YASDQAILDSGLKLTDFDLLOALLOSVAANNKAELLKEMQNDPVPVPGKTPAIAQSLV 239

Db 167 LEKQNTLKTTLTTDSADQIPAINSOLEINKNSADQIIKLEQNTI---SYEAVLTNAG 223

QY 240 DQTDATATQTEKQGNALRDAYFAGQNASGAVENAKSNNSISNDSAKAATATAKTQTAEA 299

Db 224 EVIKASSEAGIKLQALQSLVDAGDQSQAAVLQAOQNNSPDNTAATKLLDAETKYNEL 283

QY 300 QKK---PPDSPILQEAQEMVIAEQKLNKIKPADGSDVPNPGTTVGGSKQGGSSIGSI-- 354

Db 284 KOEHTGLTDSPLVKKAPEQISQAQKQTEIKPS-GSDIPTVGFPS-GSAAGSAGAVGALKS 341

QY 355 -----RVSMLLDDAENETASILMSGFRQMTLHMTENTPDSQAQQOELAAQARAAGDSDSAAALADAQKALEALGK 406

Db 342 SNNSGRISLLELDDVDNEMAAIAMQGFMSMEIQENNNPATAKELQAMEAQLTAMSDQLVG 401

QY 407 GDSAAALADAQKALEALGKAGQGGIILNALGQIASAAVVSAGV 451

Db 402 ADGELPAEIQAIKDALAAL-KQPSDTGLATAMQOVAFRAAKVGG 445

RESULT 11

ADD43866

ID ADD43866 standard; protein; 647 AA.

XX AC ADD43866;

XX 15-JAN-2004 (first entry)

XX Chlamydia trachomatis immunogenic protein, SEQ ID NO 161.

KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
 XX gene therapy; antibacterial.

OS Chlamydia trachomatis.

XX WO2003049762-A2.

XX 19-JUN-2003.

XX 12-DEC-2002; 2002WO-IB005761.

XX 12-DEC-2001; 2001GB-00029732.

XX 06-AUG-2002; 2002GB-00018233.

XX 14-AUG-2002; 2002GB-00018924.

XX (CHIR-) CHIRON SPA.

XX Grandi G, Ratti G;

XX WPI; 2003-532882/50.

XX N-PSDB; ADD43867.

XX New immunogenic composition having a protein or encoding nucleic acid,

XX useful for diagnosing, preventing and/or treating Chlamydia trachomatis

XX infection.

XX Claim 6; SEQ ID NO 161; 164pp; English.

XX The invention relates to a novel immunogenic composition comprising a
 CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
 CC acid comprises any of 131 fully defined amino acid or nucleotide
 CC sequences given in the specification, or has 50% or greater sequence
 CC identity to it, or their fragments. The protein and/or nucleic acid of
 CC the immunogenic composition is useful in the manufacture of a medicament
 CC for the treatment or prevention of infection due to Chlamydia
 CC trachomatis. The infection is treated or prevented by the medicament
 CC eliciting an immune response which is specific to a C. trachomatis
 CC elementary body, or for neutralising C. trachomatis elementary bodies,
 CC hence the immunogenic composition can be used in creating a vaccine. The
 CC immunogenic compositions can also be used for the diagnosis of C.
 CC trachomatis infection. The nucleic acids of the immunogenic compositions
 CC can be used to treat disorders by gene therapy. The immunogenic
 CC compositions have antibacterial activity. This sequence represents one of
 CC the 131 C. trachomatis proteins with immunogenic properties of the
 CC invention.

XX Sequence 647 AA;

Query Match 20.9%; Score 461.5; DB 7; Length 647;

Best Local Similarity 29.7%; Pred. No. 7.9e-24;

Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;

Qy 2 VNPFGPIDEPTERTPPADLSAQGLEASANKSAEAOIRIAGAEAKPKESKTDVSVERWSIL 61

Db 15 MNPIINGQI-----ASNSETKESKSEA-----SPSASSSVSSNSFL 52

Qy 62 RSAYNALMSLADKLGIIASSNSSSTSRG-ADVDSTTATAPPPPTDDYKQTQATYDT 120

Db 53 SSAXHALISLRD--AILNKNSSPTDLSLQLEASTSTSTVTRVAARDYNEAKSNPDPTAKSG 110

Qy 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEE-TAJAAEWETKNADAVKGAQITELAK 179

Db 111 LENATTIAEYTKWADLMWALQDMERLAKQAEVTRKEALQEQ-----EVIDKLNQLVR 166

Qy 180 YASDNQAILDSLGLTGFDDLQALLOSVAANNKAAELLKEMQDNPPVPGKTPAIAQSLV 239

Db 167 LEKQNQLTKETLTITDSADQIPAINSQLINKNSADQIILKLEQNI---SYEAVLTNAG 223

Qy 240 DQATATQIEKGNATRDYFAGQNASGAVENAKSNNSINIDSAKAAATATAKTQIAEA 299

Db 224 EVIKASEAGIKLQALQSIYDADGQSQAAVLQAGQNNSPDNIAATKKLIDAAETKYNEL 283

Qy 300 QKK---FPDSPILQAEQWVIAEKDLKNIKPADGSDVPNPGTTVGSKQGGSSIGSI-- 354

Db 284 KQERTGLTSDPLVKKAEQISQAKDIQEIKPSS-GSDIPIVGPS-GSAAGSAGVAGLKS 341
 Qy 355 -----RVSMLLDDAENETASILMSGRQMTHMENTENPDSSAAQOELAAQARAAK---AA 406
 Db 342 SNNSGRISLLDDVDNEMALAMQGRSMIEQFNVNPNPATAKEIQAEQAQTMSDQIVG 401
 Qy 407 GDDSAALADAAQKALEAALGKAGQQQGIILNALGQIASAAVVSAG 451
 Db 402 ADGELPAEIQAIDALAAL-KQPSSTDGLATAMGQVAFAAKVG 445

RESULT 12

AAU38921

ID AAU38921 standard; protein; 660 AA.

AC AAU38921;

DT 11-SEP-2003 (revised)

DT 16-JAN-2002 (first entry)

XX C. trachomatis CT87s protein.

XX Chlamydia: sexually transmitted disease; PID; antibacterial;

XX pelvic inflammatory disease; antigen; trachoma; gynecological;

XX acute respiratory tract infection; atherosclerosis; male infertility;

XX coronary heart disease.

XX Chlamydia trachomatis; serovar E.

XX WO200181379-A2.

XX 01-NOV-2001.

XX 23-APR-2001; 2001WO-US013081.

XX 21-APR-2000; 2000US-0198853P.

XX 20-JUL-2000; 2000US-0219752P.

XX (CORI-) CORIXA CORP.

XX Bhatia A, Probst P, Stromberg EJ;

XX WPI; 2001-616771/71.

XX N-PSDB; AAS56995.

XX New polynucleotide for treating Chlamydia infections encodes a
 XX polynucleotides containing an immunogenic portion of a Chlamydia antigen.

XX Claim 3; Page 205-207; 208pp; English.

XX The invention relates to isolated polynucleotide encoding at least a
 CC partial Chlamydia protein which is an antigenic fragment, or the
 CC complements, fragments, homologues and variants, and antibodies raised
 CC against the antigenic proteins (or fragments). The nucleic acids,
 CC proteins and antibodies are used to diagnose and treat Chlamydia
 CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
 CC disease (PID), acute respiratory tract infection, trachoma,
 CC atherosclerosis and coronary heart disease) in a patient, and in the
 CC treatment of male infertility. The compounds of the invention are also
 CC useful for detecting the presence of Chlamydia in a patient, and
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-
 CC 2003 to standardise OS field)

XX Sequence 660 AA;

Query Match 20.8%; Score 453; DB 4; Length 660;

Best Local Similarity 28.9%; Pred. No. 3.2e-23;

Matches 136; Conservative 94; Mismatches 193; Indels 48; Gaps 14;

Qy 6 GPGPIDERTTPADLSAQGLEASANKSAEAOIRIAGAEAKPKESKTDVSVERWSILRSV 65

11	DB	GPESVSNQSSMNP	IIINGQIASNSETKSTKA-----SEASP--SASSSVSSWFLSSAK	63
66	QY	NALMSIADKLGIASNS	SSSSSR--ADVDSNTATAPPTPPTDDYKTAQTHAYDTFTS	124
64	DB	NALISLRD--AL	LNKSSPTLSOLEASTSTSTVVAADYDEAKSNFDTAKSGLENA	121
125	QY	TSJALQIAALVSDQAV	NIKDTAATD-----EET-----AJAAEWETKNADAVKVGQ	173
122	DB	KTLASVETKMDLMA	LQDMERLANDSPSNHTEVNIKALEAQDKTID-----K	173
174	QY	ITELAKIASDNOA	ILDSLGKLTFTDLLQAAALQSVANNKAAELKEMQDNPVPGKTPA	233
174	DB	LKNLVTLQONKSL	TEVLKTTSDADQIPAINSOLEINKNSADQLIKDLERQNI---SYFA	230
234	QY	IAGSIVDQCDATAT	OIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAALATAK	293
231	DB	VLTNAGEVIKASSE	AGIKGLQALQISVDAGQSQAAVLQAQONNSPDNTAATKELIDAE	290
294	QY	TQIAERAQK---FP	SPPILOEABQNVIOAEKDLKNIIPKADGSDVPNPNGTTVGSGKQGS	350
291	DB	TKVNELKQEH	TGLTDSPLVKXAEQISQAQXDIQEIKPS--GSDIPVGPS--GSAASAGSA	348
351	QY	IGSI-----RV	SMLDDAENETASILMSGFROMIHFNTENPDSQAQOELAAQARA	403
349	DB	AGALKSNNSGR	SILLLDQVDNEMAALALQGRFSMIEQFNVNPNPATAKELQAMEAQLTAM	408
404	QY	K---AAGD	SAAAALADRAQALEAALGKAGQOQOQILNALGQIASAAVVSAG	451
409	DB	SDQLVGADGEL	FAETQATKDALAQAL--KQPSADGLATAMQGVAFAPAAKVG	458

RESULT 13
ADD42726
ID ADD42726 standard; protein; 660 AA.
XX AC
XX ADD42726;
XX 15-JAN-2004 (first entry)
XX DT
XX DE Chlamydia trachomatis antigen protein SEQ ID NO:139.

The present invention describes compounds and methods for diagnosing and treating Chlamydial infection. Chlamydia polynucleotide and protein sequences have antibiotic, antiinflammatory, antifertility, cardiant, antiarteriosclerotic and ophthalmological activities, and can be used in vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins, compositions or methods from the present invention can be used for the serodagnosis or treatment of Chlamydial infections, particularly in humans. The polynucleotides, proteins or compositions are particularly useful for stimulating an immune response in a patient, or for stimulating and/or expanding T cells specific for a Chlamydia protein. Specifically, the polynucleotides, proteins or compositions are useful as vaccines for treating or preventing Chlamydial infections including pelvic inflammatory disease (which results in tubal obstruction and infertility in women), male infertility, ocular infection (which may cause blindness), acute respiratory tract infections, atherosclerosis, or coronary heart disease. The present sequence is used in the exemplification of the present invention.

RESULT 14	
AA37571	
ID	AA37571 standard; protein, 350 AA.
XX	
XX	
AC	AA37571;
XX	
XX	
DT	07-OCT-1999 (first entry)
XX	
XX	
DE	Chlamydia trachomatis surface exposed protein.
XX	
XX	
KW	Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW	paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW	nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW	bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX	
XX	

```
OS Chlamydia trachomatis.
XX WO928475-A2.
XX 10-JUN-1999.
XX 27-NOV-1998; 98WO-18001939.
XX 28-NOV-1997; 97FR-00015041.
XX 17-DEC-1997; 97FR-00015034.
XX 04-NOV-1998; 98US-0107077P.
XX (GEST ) GENSET.
XX Griffais R;
XX WPI; 1999-371125/31.
XX Genome sequence of Chlamydia trachomatis.
XX Disclosure; Page 1226-1227; 1755pp; English.
XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
XX of Chlamydia trachomatis (see AYZ01425). The polypeptides can be used as
XX vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
XX can also be used to control growth of the microorganism. Chlamydia
XX trachomatis is responsible for a large number of diseases, e.g. eye
XX diseases such as conventional trachoma, nonendemic trachoma,
XX paratrachoma, and inclusion conjunctivitis; genital diseases such as
XX nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
XX perinephritis, Bartholinitis; pneumopathy in breast feeding infants; and
XX venereal lymphogranulomatosis. The polypeptides of the invention may be
XX of use in treating these diseases
XX Sequence 350 AA;

Query Match 13.1%; Score 289.5; DB 2; Length 350;
Best Local Similarity 24.9%; Pred. No. 4.1e-12;
Matches 94; Conservative 68; Mismatches 129; Indels 87; Gaps 11;

QY 2 VNPIGPGIDETERTPPADLSAQGLEASANKSAAQRIAGAEKPKESKTDVSVERWSIL 61
Db 23 MNPINGQI-----ANSETKESKASEA-----SSASSSVSSWSFL 60
QY 62 RSANALMSIADKLGIASNSSSTGRS-ADVDSTTATPTPPPTFDYKTOAQATYDT 120
Db 61 SSAKNALISLRD--AILNKNSSPTDSLQLEASTSTSTVTRVAAKYDRAKSNFTAKSG 118
QY 121 IFTS-----TSLADIQALVSLQ-----DAVTNIKDTAATDBETALAAEWETK 163
Db 119 LENAKTLAEYETKADLMALQDMEANSOPSDNHTBELNNIKALEAQKDTI----- 170
QY 164 NADAVKVGAGITELAKYASDNQAILSLGKLTGFDLLQALQSVANNKKAELLKEMOD 223
Db 171 -----DKLNKLVILQNKQSLTEALTKTDSADQIPAINSLINKNSAHQIILKE 222
QY 224 NFVVPKTPALQSLVDQDTATATQTEKO-----GNAIRDAYFAGQNASGAVEN 272
Db 223 Q-----ISNYKAVLTIDVEKVIKEFSEAGIKGLQALQSLVDAGDQSQAAVLQ 268
QY 273 AKSNNSISNTDSAKALATAKTQIAEAQ---KKFPDSPILOEAEOMVIOAEKDLKNKEA 329
Db 269 ARGNSPNDIAAATKILDAKTKNELKQBHEIADSPLVKKAEBQINQAQDQIQTITPS 328
QY 330 DGSDFVNPPTGTVGSGSKQ 347
Db 329 -GLDIPIVGSGSGXPQE 345

RESULT 15
AAB13695
ID AAB13695 standard; protein; 361 AA.
XX
```

```
AC AAB13695;
XX
XX DT 02-FEB-2001 (first entry)
XX
XX DE Chlamydia sp. protein # 6.
XX
XX KW Chlamydial infection; sexually transmitted disease;
XX KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
XX KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;
XX KW coronary heart disease; antibacterial.
XX
XX OS Chlamydia sp.
XX
XX PN WO200034483-A2.
XX
XX PD 15-JUN-2000.
XX
XX PF 08-DEC-1999; 99WO-US029012.
XX
XX PR 08-DEC-1998; 98US-00208277.
XX PR 08-APR-1999; 99US-00288594.
XX PR 01-OCT-1999; 98US-00410568.
XX PR 22-OCT-1999; 99US-00426571.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX WPI; 2000-431303/37.
XX
XX DR Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX PT comprises immunogenic portion of Chlamydia antigen, which comprises amino
XX PT acid sequence encoded by polynucleotide sequence.
XX
XX PS Disclosure; Page 251-253; 256pp; English.
XX
XX CC The present invention relates to new nucleic acid sequences and the
XX CC proteins encoded by the nucleic acid sequences. The encoded proteins
XX CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX CC proteins are useful for the serodiagnosis and treatment of Chlamydia
XX CC infection. Chlamidiae are intracellular bacterial pathogens that are
XX CC responsible for a wide variety of human infections. C. trachomatis
XX CC infection is one of the most common sexually transmitted diseases and can
XX CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX CC and infertility. Trachoma due to ocular infection with C. trachomatis is
XX CC the leading cause of preventable blindness worldwide. C. pneumonia is a
XX CC major cause of acute respiratory tract infections in humans and is also
XX CC thought to play a role in the pathogenesis of atherosclerosis and
XX CC coronary heart disease. The present sequence is a protein isolated in the
XX CC present invention
XX
XX SQ Sequence 361 AA;

Query Match 10.5%; Score 231.5; DB 3; Length 361;
Best Local Similarity 35.4%; Pred. No. 5e-08;
Matches 57; Conservative 30; Mismatches 61; Indels 13; Gaps 5;

QY 301 KKFPDSPILOEAEOMVIOAEKDLKNKPADGSDVNPPTGTVGSGKQGGSSIGSI----- 354
Db 2 QBIADSPLVKKAEBQINQAQDQIQTITPS-GLDIPIVGPS-GSAASASAGALSSNNS 59
QY 355 -RVSMLLDDAENETASILMSGFRQMHMTENTENPDSQAQOELAAQAPAAK---AAGDSD 410
Db 60 GRISLLDDVDNEMAIAIAQGFMSIEQFNVNPNATAKELQAMEAQLTAMSDQLVGADGE 119
QY 411 ARAALADQAQKALEALGKAGQQQGLINLALGQIASNAVVSAG 451
Db 120 LPAEIQATKDALAQAL-KQPSADGLATAMGQVFAAAXVGG 159

Search completed: March 23, 2004, 19:27:40
Job time : 40.1412 secs
```


QY 421 ALEAALGKAGQQQGGILNALGQIASAAV 447
Db 423 ALEAALGKAGQQQGGILNALGQIASAAV 449

RESULT 2
US-09-556-877-299
; Sequence 299, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-299

Query Match 10.5%; Score 231.5; DB 4; Length 361;
Best Local Similarity 35.4%; Pred. No. 1e-09;
Matches 57; Conservative 30; Mismatches 61; Indels 13; Gaps 5;

QY 301 KKPDPSPILQAEQMVQIAEKDLKNKPADGSDVPNPGTTVGSKQCGSSIGSI----- 354
Db 2 QEIADSPLVKAEQINQAQDIQTTPS-GLDPIVGPS-GSAASAGSAGALKSSNNS 59

QY 355 -RVSMILLDDAENETASILMSGFRQMHMTENTPDSQAQQELAAQAAK---AAGDSDS 410
Db 60 GRISLLDDVDNEMAAIAMQGFMSIEQFNVPNPATAKELQAMEAQLTAMSDQLVGADGE 119

QY 411 AAAALADAQKALEAALGKAGQQQGGILNALGQIASAAVVSAG 451
Db 120 LPAEIQAIKDALAQAL-KQPSADGLATAMQVAFAAAKVGG 159

RESULT 3
US-09-620-412C-299
; Sequence 299, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-299

Query Match 10.5%; Score 231.5; DB 4; Length 361;
Best Local Similarity 35.4%; Pred. No. 1e-09;
Matches 57; Conservative 30; Mismatches 61; Indels 13; Gaps 5;

QY 301 KKPDPSPILQAEQMVQIAEKDLKNKPADGSDVPNPGTTVGSKQCGSSIGSI----- 354
Db 2 QEIADSPLVKAEQINQAQDIQTTPS-GLDPIVGPS-GSAASAGSAGALKSSNNS 59

QY 355 -RVSMILLDDAENETASILMSGFRQMHMTENTPDSQAQQELAAQAAK---AAGDSDS 410
Db 60 GRISLLDDVDNEMAAIAMQGFMSIEQFNVPNPATAKELQAMEAQLTAMSDQLVGADGE 119

QY 411 AAAALADAQKALEAALGKAGQQQGGILNALGQIASAAVVSAG 451
Db 120 LPAEIQAIKDALAQAL-KQPSADGLATAMQVAFAAAKVGG 159

RESULT 4
US-09-598-419-299
; Sequence 299, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-299

Query Match 10.5%; Score 231.5; DB 4; Length 361;
Best Local Similarity 35.4%; Pred. No. 1e-09;
Matches 57; Conservative 30; Mismatches 61; Indels 13; Gaps 5;

QY 301 KKPDPSPILQAEQMVQIAEKDLKNKPADGSDVPNPGTTVGSKQCGSSIGSI----- 354
Db 2 QEIADSPLVKAEQINQAQDIQTTPS-GLDPIVGPS-GSAASAGSAGALKSSNNS 59

QY 355 -RVSMILLDDAENETASILMSGFRQMHMTENTPDSQAQQELAAQAAK---AAGDSDS 410
Db 60 GRISLLDDVDNEMAAIAMQGFMSIEQFNVPNPATAKELQAMEAQLTAMSDQLVGADGE 119

QY 411 AAAALADAQKALEAALGKAGQQQGGILNALGQIASAAVVSAG 451
Db 120 LPAEIQAIKDALAQAL-KQPSADGLATAMQVAFAAAKVGG 159

RESULT 5
US-09-134-001C-4463
; Sequence 4463, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4463
; LENGTH: 2137
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4463

Query Match 8.2%; Score 180; DB 4; Length 2137;
Best Local Similarity 19.7%; Pred. No. 9.1e-05;
Matches 84; Conservative 89; Mismatches 184; Indels 70; Gaps 13;

QY 19 ADLSAQGLEASAAKSAEAQRIAGAEAKPKESKTDSVERWSITRSVAVNALMSLADKLGIA 78

Db 1155 ASTSLGSLSTISDSTSTSTSDASTSTSESDSTST-SLSBSTST---SLSDSTST 1210
Qy 79 SSSSS-----STSRADVDSTATAPPPPTFDYKTAQATYDTTFTSTSLADIQ 131
Db 1211 TSESASTSTSESDSTSESTSTSTSVSD-----STSASTSDASTSTSVSDSE 1261
Qy 132 AALVSLQAV-TNKDRTAATDEETAAAEWETKNADAVKVGQAITE-----176
Db 1262 SASSTISGSLSTSVSDSTST--STSDASTSTSESDSTSESTSTSESTSTSVSDSTGAST 1319
Qy 177 -----LAKYASDNQAILDGLKLTSTFLLQALLOSANNKAELKEMQDNVPVFGKT 231
Db 1320 SDSASTSTSESDASTSTSGSTSTSLSDST---STSDASTSTSESDSEASTSLS 1376
Qy 232 PAIAQSLVDQDTATATOTEKDGNAIRDAYFAGNASGAVENAKNN-SISNIDSAKAAIA 290
Db 1377 GSTSTSLSDSTSTSDASTSTSVSDSASTSTSLGSLSTSVSDSTSTSDSASTS 1436
Qy 291 TAKTQIAEAAQKFPDPSILQAEQWVQAEKDLKNIKPADGSDVNPPTTVGGSKQGGSS 350
Db 1437 ESDSERASTSLGSTSTISDSTST-----TSTDASTSTSTSVSESNSTST 1482
Qy 351 IG---SIRVSMLLDAAENETASILMSGFROMIHMENTENPDSSQAQOELAAQAAKAG 407
Db 1483 ISESLSTSVSDSTSTSTSDASTSTST-----VSDSDASTSTSESVSDSESTSTST 1535
Qy 408 DDSAAAA 414
Db 1536 SDSASTS 1542

RESULT 6
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match 7.4%; Score 163.5; DB 4; Length 10182;
Best Local Similarity 21.8%; Pred. No. 0.014;
Matches 108; Conservative 66; Mismatches 194; Indels 127; Gaps 20;
Qy 16 TPPADLSAQGLEASAANKSAEQAQRIAGAE-----KPKES-----KTDVS 55
Db 2628 TPQVNSALSKVQAQNKINEAKALLQNKADNSQLVRAKEQLQSQIQPAASTDGMTQDST 2687
Qy 56 ERWSILSAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDY-----110
Db 2688 RNYKNKRAEQAQIQAHSV---INNGDATSQINDAKNTVEQAQR-----DYVEAKS 2737
Qy 111 -----KTAQATYDTTFTSTSLADIQAALV-SLODAVTNIKDTATDEETAAAEWETKN 164
Db 2738 NLRADKSLQAYDTLNRDVLTDNKKPASVRRYNEAISNIR-----KE 2780
Qy 165 ADAVKVGAQITELAKYASDNQAILDGLKLTSTF--LLOA-ALLQSVANNKAELKEM 221
Db 2781 LDTAKADAST-LANTNPSVEQVRDALNKINTVPKVNQAIALQPKENNELVQAKRL 2839

Qy 222 QDN-----PVVPGKTP-----AIAQSLVDQDTATATQIEKDGNAIRDAY 260
Db 2840 QDAVNDIPQGTQGTQTTNNYNDKQREARALTSARVINDGDTTQETITSEKSKVEQAM 2899
Qy 261 FAGNASG-----AVENAKNN-----SISNIDSAKAAIATAKQIAEAAQK 301
Db 2900 QALTNAKSNLRADKNELQATYAKNLIENVSTNGKKPASIRQYETAKARI---QNGINDAKN 2956
Qy 302 KFPDPSILQAEQWVQAEKDLKNIKPADGSDVNPPTTVGGSKQGGSSIRVSMLLD 361
Db 2957 E--AERILGNDNPQSVTQALNKIKAIQ---PKLTEAINMLQNKENNELVNAKNRLE 3010
Qy 362 DAENETASILMSGFROMIHMENTENPDSSQAQOEL-----AAQARAAKAAGDD 409
Db 3011 NAVNDTDP--THGWTQ--ETINNYNKAKREAQNEIQKANMIINNGDATAQDISSEKSKVE 3066
Qy 410 SAAALADAQKALEA 424
Db 3067 QVLQALQNAKNDLEA 3081

RESULT 7
US-08-216-894-8
; Sequence 8, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchhoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPAOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-216-894-8

Query Match 7.0%; Score 154.5; DB 2; Length 643;
Best Local Similarity 21.5%; Pred. No. 0.0014;
Matches 99; Conservative 59; Mismatches 168; Indels 135; Gaps 17;
Qy 7 PGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEKPKSKTDSVERWSILSAVN 66
Db 216 PPKSDILVPRGSPQL--QQAENNTITSKKEMTKLREKVKKAKEKLDAINRATKLEERN 273
Qy 67 ALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATYDTTST 126
Db 274 QAYKAHK---ACEEKAKTQRLITPESINLNKKRP-----307

Query Match 7.0%; Score 154.5; DB 3; Length 643;
Best Local Similarity 21.5%; Pred. No. 0.0014;
Matches 99; Conservative 59; Mismatches 168; Indels 135; Gaps 17;

RESULT 9


```

QY 311 -EAEQWVIOAEKOLKNTKPADGSDVPNPGTTVGSGKQGSIGSIRVSMLLDDAENETAS 369
Db 285 ADYEAELQAEYKDLAAQSGSNATNEADYQAKAAVEQELARV-----QAANAAG 334
QY 370 ILMSGFQMHMENTENPDSQAQOELAAQARAARAAAGDPSAAALADAAQALEAA 425
Db 335 ---QAYEQALAAANTAKN--AQITAEAIQORNAQAKA--NYEAKLQAYOKDLAAA 383

RESULT 14
US-08-216-894-2
; Sequence 2, Application US/08216894
; Patent No. 5876734
; GENERAL INFORMATION:
; APPLICANT: Kirchhoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,894
; FILING DATE: 24-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-216-894-2

Query Match 6.5%; Score 142.5; DB 2; Length 564;
Best Local Similarity 21.4%; Pred. No. 0.0095;
Matches 97; Conservative 55; Mismatches 164; Indels 137; Gaps 17;

QY 7 PGPIDETPTPPADLSAQGLEASAAKSAQAORIAAGAEAKPKESKTSVERWSILRSVAVN 66
Db 216 PPKSGLVPRGSPQL--QQAENNIITNSKKEMTKLREKVKAAEKLDAINRAKLEERN 273
QY 67 ALMSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTDDYKTOAQATYDTFTSTS 126
Db 274 QAYKAAHK---ABEEKAKTFQRLITFSEINILKRP----- 307
QY 127 LADIQAAALVSLQAVTN-----IKDTATDE-----ETAIAAEWETKNADAVKVAQI 174
Db 308 -----NDVSNRDKKNSSETAKTDEVEKQRAAEAAKAVETEKORA-----AEA 350
QY 175 TELAKYASDNQAILDSGLKLTSLDQLAALLQSVANNKAAELLKEMQDNVVPVPGKTPAI 234
Db 351 TKVAE-----AEKKAAAEAAKAVETEKORAAEATKV 381
QY 235 A----QSLVQDTATATQIEKGNARIDAVFAGNAGAVENAKSNISNIDSAKAAIA 280
Db 382 AEAEKQKAAAEAAKAVETEKORAAEATKVAAEAKQRAAEAMK-----VAEAEKQKAAEA 434

```

```

QY 291 TAKTOIAEAKK-KFPDSPILQEAQM-----VIOAEKDLKNIKPADGSDVPNPGTTVG 342
Db 435 ---TKVAEAEKQKAAEATKVAAEAKQKAAEATKVAAEAKQ---KAAEATKVAE----- 481
QY 343 GSKQOQSSIGSIRVSMMLDDAENETASILMSGFQMHMENTENPDSQAQOELAAQARA 402
Db 482 AEKQKAAEATKVA-----EAEKQKAAEATK-----VAEAEKQKAAEATKVAAEAK 526
QY 403 AKAGDSDSAAAA---LADAQKALEAALGKAGQ 431
Db 527 QKAAEATKVAAEAKQKAAEATKVAAEAKQKAG 559

RESULT 15
US-09-115-746-2
; Sequence 2, Application US/09115746
; Patent No. 6228601
; GENERAL INFORMATION:
; APPLICANT: Kirchhoff, Louis V.
; APPLICANT: Otsu, Keiko
; TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
; TITLE OF INVENTION: WITH TRYPANOSOMA CRUZI
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/115,746
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,894
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 85326/102/DRLO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-115-746-2

Query Match 6.5%; Score 142.5; DB 3; Length 564;
Best Local Similarity 21.4%; Pred. No. 0.0095;
Matches 97; Conservative 55; Mismatches 164; Indels 137; Gaps 17;

QY 7 PGPIDETPTPPADLSAQGLEASAAKSAQAORIAAGAEAKPKESKTSVERWSILRSVAVN 66
Db 216 PPKSGLVPRGSPQL--QQAENNIITNSKKEMTKLREKVKAAEKLDAINRAKLEERN 273
QY 67 ALMSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTDDYKTOAQATYDTFTSTS 126
Db 274 QAYKAAHK---ABEEKAKTFQRLITFSEINILKRP----- 307
QY 127 LADIQAAALVSLQAVTN-----IKDTATDE-----ETAIAAEWETKNADAVKVAQI 174
Db 308 -----NDVSNRDKKNSSETAKTDEVEKQRAAEAAKAVETEKORA-----AEA 350

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 05:59:21 ; Search time 105.628 Seconds
(without alignments)
1108.119 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204

Sequence: 1 MWNPIGPGPIDETERTPPAD.....QGILNALGQIAAAVVSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/2/pubpa/US06_PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubpa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubpa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpa/US09_PUBCOMB.pap.*
10: /cgn2_6/ptodata/2/pubpa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpa/US09_PUBCOMB.pap.*
12: /cgn2_6/ptodata/2/pubpa/US09_NEW_PUB.pap.*
13: /cgn2_6/ptodata/2/pubpa/US10_PUBCOMB.pap.*
14: /cgn2_6/ptodata/2/pubpa/US10_PUBCOMB.pap.*
15: /cgn2_6/ptodata/2/pubpa/US10_PUBCOMB.pap.*
16: /cgn2_6/ptodata/2/pubpa/US10_NEW_PUB.pap.*
17: /cgn2_6/ptodata/2/pubpa/US60_NEW_PUB.pap.*
18: /cgn2_6/ptodata/2/pubpa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2204	100.0	452	10	US-09-564-479-6
2	2204	100.0	490	10	US-09-564-479-8
3	2204	100.0	651	9	US-09-841-132-405
4	2204	100.0	651	10	US-09-564-479-2
5	2204	100.0	651	10	US-09-841-260-78
6	2204	100.0	651	13	US-10-007-693-78
7	2204	100.0	651	15	US-10-312-273-13
8	2182	99.0	478	15	US-10-289-762-776
9	1866	84.7	583	10	US-09-564-479-4
10	461.5	20.9	647	9	US-09-841-132-436
11	453	20.6	660	10	US-09-841-260-139
12	453	20.6	660	13	US-10-007-693-139
13	231.5	10.5	361	9	US-09-841-132-239
14	219	9.9	1463	9	US-09-971-536-69
15	190	8.6	2478	9	US-09-815-242-5816

16 190 8.6 2478 9 US-09-815-242-12967 Sequence 12967, A
17 189.5 8.6 1831 12 US-10-282-122A-71033 Sequence 71033, A
18 188 8.5 603 12 US-10-282-122A-72004 Sequence 72004, A
19 186.5 8.5 790 12 US-10-282-122A-75142 Sequence 75142, A
20 184 8.3 1122 12 US-10-282-122A-42783 Sequence 42783, A
21 181 8.2 2481 12 US-10-282-122A-43762 Sequence 43762, A
22 178 8.1 1178 14 US-10-128-714-8240 Sequence 8240, Ap
23 176 8.0 1924 12 US-10-282-122A-43924 Sequence 71846, A
24 173 7.8 2271 12 US-10-282-122A-43924 Sequence 43924, A
25 172.5 7.8 812 12 US-10-282-122A-75242 Sequence 75242, A
26 172 7.8 2344 9 US-09-815-242-12713 Sequence 12713, A
27 171.5 7.8 1879 9 US-09-971-536-70 Sequence 70, Appl
28 170.5 7.7 2368 9 US-09-815-242-5635 Sequence 5635, Ap
29 170.5 7.7 2368 9 US-09-815-242-12389 Sequence 12389, A
30 168.5 7.6 2398 12 US-10-282-122A-70176 Sequence 70176, A
31 168.5 7.6 6281 9 US-09-815-242-12996 Sequence 12996, A
32 168 7.6 993 12 US-10-282-122A-43875 Sequence 43875, A
33 168 7.6 2283 14 US-10-172-502-4-71885 Sequence 4, Appl
34 166.5 7.6 469 12 US-10-282-122A-70453 Sequence 70453, A
35 166.5 7.6 1665 12 US-10-282-122A-71690 Sequence 71690, A
36 166 7.5 2086 9 US-09-815-242-5639 Sequence 5639, Ap
37 166 7.5 5795 9 US-09-815-242-12610 Sequence 12610, A
38 165.5 7.5 385 12 US-10-282-122A-71885 Sequence 71885, A
39 165.5 7.5 454 12 US-10-424-599-232134 Sequence 232134, A
40 164 7.4 6713 12 US-10-282-122A-43811 Sequence 43811, A
41 163.5 7.4 1123 12 US-10-282-122A-70581 Sequence 70581, A
42 163 7.4 1348 12 US-10-282-122A-56877 Sequence 56877, A
43 162.5 7.4 2437 9 US-09-815-242-5834 Sequence 5834, Ap
44 162 7.4 722 9 US-09-815-242-12888 Sequence 12888, A
45 162 7.4 991 9 US-09-815-242-5803 Sequence 5803, Ap

ALIGNMENTS

RESULT 1
US-09-564-479-6
; Sequence 6, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-6

Query Match 100.0%; Score 2204; DB 10; Length 452;
Best Local Similarity 100.0%; Pred. No. 2.9e-160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEORAGAEAKPKSKTDSVERWSI 60
Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEORAGAEAKPKSKTDSVERWSI 60
QY 61 LRSVAVNALMSLADKLGIASSNSSSSSRSADVDSITATPTPPPTFDYKTAQAYDT 120
Db 61 LRSVAVNALMSLADKLGIASSNSSSSSRSADVDSITATPTPPPTFDYKTAQAYDT 120

121 IFSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
121 IFSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNPVVPKTPAIQSLVD 240
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNPVVPKTPAIQSLVD 240
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQOQSSIGSIRVSMML 360
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQOQSSIGSIRVSMML 360
361 DDAENETASILMSGFRQMHMENTENPDQSAQAQELAAQARAARAAAGDDSAALADAQK 420
361 DDAENETASILMSGFRQMHMENTENPDQSAQAQELAAQARAARAAAGDDSAALADAQK 420
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 2

US-09-564-479-8
; Sequence 8, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: COHEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; FILE REFERENCE: US3 THEREOF

; CURRENT APPLICATION NUMBER: US/09/564,479

; CURRENT FILING DATE: 2000-05-03

; PRIOR APPLICATION NUMBER: 60/132,270

; PRIOR FILING DATE: 1999-05-03

; PRIOR APPLICATION NUMBER: 60/141,276

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 8

; LENGTH: 490

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-564-479-8

Query Match 100.0%; Score 2204; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.2e-160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDDYKTAQATAYDT 120
61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDDYKTAQATAYDT 120
121 IFSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
121 IFSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNPVVPKTPAIQSLVD 240
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNPVVPKTPAIQSLVD 240
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300

241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQOQSSIGSIRVSMML 360
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQOQSSIGSIRVSMML 360
361 DDAENETASILMSGFRQMHMENTENPDQSAQAQELAAQARAARAAAGDDSAALADAQK 420
361 DDAENETASILMSGFRQMHMENTENPDQSAQAQELAAQARAARAAAGDDSAALADAQK 420
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 3

US-09-841-132-405
; Sequence 405, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; FILE REFERENCE: DIAGNOSIS OF CHLAMYDIAL INFECTION

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 405

; LENGTH: 651

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-841-132-405

Query Match 100.0%; Score 2204; DB 9; Length 651;
Best Local Similarity 100.0%; Pred. No. 4.8e-160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEQAQIAGAEAKPKESKTDTSVERWSI 60
61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDDYKTAQATAYDT 120
61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDTTATPTPPPTFFDDYKTAQATAYDT 120
121 IFSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
121 IFSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNPVVPKTPAIQSLVD 240
181 ASDNQAILDSLGLKTSFDLLQALLOSVAANNKAAELLKEMQNPVVPKTPAIQSLVD 240
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
241 QTDATATQIEKGNADIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQ 300
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQOQSSIGSIRVSMML 360
301 KFPDPSPILOEAEQWVIOAEKDLKNIKPADGSDVPNPPTTVGGSKQOQSSIGSIRVSMML 360
361 DDAENETASILMSGFRQMHMENTENPDQSAQAQELAAQARAARAAAGDDSAALADAQK 420
361 DDAENETASILMSGFRQMHMENTENPDQSAQAQELAAQARAARAAAGDDSAALADAQK 420
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 4

US-09-564-479-2
 ; Sequence 2, Application US/09564479
 ; Publication No. US20030095973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: COMEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; CURRENT FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR APPLICATION NUMBER: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 ; US-09-564-479-2

Query Match 100.0%; Score 2204; DB 10; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.8e-160;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAANKSABQAQRIAGAEAKPKESKTSVERWSI 60
 DB 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAANKSABQAQRIAGAEAKPKESKTSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 QY 181 ASDNQAILDSGLKTSFDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 DB 181 ASDNQAILDSGLKTSFDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPPDSPILQEAQMVQAEKDLKNIKPADGSDVPNPFGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPPDSPILQEAQMVQAEKDLKNIKPADGSDVPNPFGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 5

US-09-841-260-78
 ; Sequence 78, Application US/09841260
 ; Publication No. US20030175700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Stromberg, Erika Jean
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
 ; OF CHLAMYDIAL INFECTION

FILE REFERENCE: 210121.515
 ; CURRENT APPLICATION NUMBER: US/09/841,260
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 140
 ; SEQ ID NO 78
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 ; US-09-841-260-78

Query Match 100.0%; Score 2204; DB 10; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.8e-160;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAANKSABQAQRIAGAEAKPKESKTSVERWSI 60
 DB 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAANKSABQAQRIAGAEAKPKESKTSVERWSI 60
 QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVDSSTTATPTPPPTFDDYKTAQTAYDT 120
 QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 DB 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 QY 181 ASDNQAILDSGLKTSFDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 DB 181 ASDNQAILDSGLKTSFDLLQAALLQSVANNKAAELLKEMQDNVVPVPGKTPAIAQSLVD 240
 QY 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 DB 241 QTDATATQIEKGNNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 QY 301 KKFPPDSPILQEAQMVQAEKDLKNIKPADGSDVPNPFGTTVGGSKQGGSSIGSIRVSMML 360
 DB 301 KKFPPDSPILQEAQMVQAEKDLKNIKPADGSDVPNPFGTTVGGSKQGGSSIGSIRVSMML 360
 QY 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 DB 361 DDAENETASILMSGFRQMIHMFNTENPDSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
 QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452
 DB 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGV 452

RESULT 6

US-10-007-693-78
 ; Sequence 78, Application US/10007693
 ; Publication No. US20020146776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; OF INFECTION/ AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515C2
 ; CURRENT APPLICATION NUMBER: US/10/007,693
 ; CURRENT FILING DATE: 2001-12-05
 ; NUMBER OF SEQ ID NOS: 157
 ; SEQ ID NO 78
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 ; US-10-007-693-78

Query Match 100.0%; Score 2204; DB 13; Length 651;
 Best Local Similarity 100.0%; Pred. No. 4.8e-160;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAANKSABQAQRIAGAEAKPKESKTSVERWSI 60
 DB 1 MVNPIGPGPIDEETERTPPADLSAAGLEASAANKSABQAQRIAGAEAKPKESKTSVERWSI 60

QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPTPTFDYKTAQATYDT 120
DB 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPTPTFDYKTAQATYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
QY 181 ASDNQAILDSLGKLTSTFLLQALQLQSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVD 240
DB 181 ASDNQAILDSLGKLTSTFLLQALQLQSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVD 240
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSKAAIATAKTAQIAEAQ 300
DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSKAAIATAKTAQIAEAQ 300
QY 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSKQOQSSIGSIRVSMML 360
DB 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSKQOQSSIGSIRVSMML 360
QY 361 DDAENETASILMSGPROMIHMFNTENPDQAAQOELAAQARAAGDSDSAAALADAOX 420
DB 361 DDAENETASILMSGPROMIHMFNTENPDQAAQOELAAQARAAGDSDSAAALADAOX 420
QY 421 ALAALGKAGQOQOILNALGQIASAAVVSAGV 452
DB 421 ALAALGKAGQOQOILNALGQIASAAVVSAGV 452

RESULT 7

US-10-312-273-13
; Sequence 13, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SA

; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE

; FILE REFERENCE: P025035W0

; CURRENT APPLICATION NUMBER: US/10/312,273

; CURRENT FILING DATE: 2002-12-20

; PRIOR APPLICATION NUMBER: 0016363.4

; PRIOR FILING DATE: 2000-07-03

; PRIOR APPLICATION NUMBER: 0017047.2

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 0017983.8

; PRIOR FILING DATE: 2000-07-21

; PRIOR APPLICATION NUMBER: 0019368.0

; PRIOR FILING DATE: 2000-08-07

; PRIOR APPLICATION NUMBER: 0020440.4

; PRIOR FILING DATE: 2000-08-18

; PRIOR APPLICATION NUMBER: 0022583.9

; PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 0027549.5

; PRIOR FILING DATE: 2000-11-10

; PRIOR APPLICATION NUMBER: 0031706.5

; PRIOR FILING DATE: 2000-12-22

; NUMBER OF SEQ ID NOS: 664

; SOFTWARE: SeqWin99, version 1.02

; SEQ ID NO 13

; LENGTH: 651

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-10-312-273-13

Query Match 100.0%; Score 2204; DB 15; Length 651;
Best Local Similarity 100.0%; Pred. No. 4,8e-160;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQIAEAKPKESKTDVSVERWSI 60
DB 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQIAEAKPKESKTDVSVERWSI 60
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPTPTFDYKTAQATYDT 120

DB 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPTPTFDYKTAQATYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
QY 181 ASDNQAILDSLGKLTSTFLLQALQLQSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVD 240
DB 181 ASDNQAILDSLGKLTSTFLLQALQLQSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVD 240
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSKAAIATAKTAQIAEAQ 300
DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSKAAIATAKTAQIAEAQ 300
QY 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSKQOQSSIGSIRVSMML 360
DB 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSKQOQSSIGSIRVSMML 360
QY 361 DDAENETASILMSGPROMIHMFNTENPDQAAQOELAAQARAAGDSDSAAALADAOX 420
DB 361 DDAENETASILMSGPROMIHMFNTENPDQAAQOELAAQARAAGDSDSAAALADAOX 420
QY 421 ALAALGKAGQOQOILNALGQIASAAVVSAGV 452
DB 421 ALAALGKAGQOQOILNALGQIASAAVVSAGV 452

RESULT 8

US-10-289-762-776
; Sequence 776, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev

; TITLE OF INVENTION: and treatment of infection

; FILE REFERENCE: 97/40-003-999

; CURRENT APPLICATION NUMBER: US/10/289,762

; CURRENT FILING DATE: 2003-03-27

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 776

; LENGTH: 478

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-10-289-762-776

Query Match 99.0%; Score 2182; DB 15; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.5e-158;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQIAEAKPKESKTDVSVERWSI 60
DB 3 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQIAEAKPKESKTDVSVERWSI 62
QY 61 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPTPTFDYKTAQATYDT 120
DB 63 LRSVAVNLSLADKLGIASSNSSSTSRSDVDSSTATATPTPTPTFDYKTAQATYDT 122
QY 121 IFTSTSLADIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 180
DB 123 IFTSTSLADIQAALVSLQDAVNIKDTAATDEETAIAAEWETKNADAVKVGAIITELAKY 182
QY 181 ASDNQAILDSLGKLTSTFLLQALQLQSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVD 240
DB 183 ASDNQAILDSLGKLTSTFLLQALQLQSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVD 242
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSKAAIATAKTAQIAEAQ 300
DB 243 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSKAAIATAKTAQIAEAQ 302
QY 301 KKFPPSPILQEAQOMVIOAEKDLKNIKPADGSDVNPVPGTIVGSKQOQSSIGSIRVSMML 360

Db 303 KKFDPSPLOAEQWVIAEQKDLNKKIPADGSDVNPFTTGGSKQGGSSIGTSIRVSMIL 362
Qy 361 DDAENETASILMSGFRQWIMHNTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQK 420
Db 363 DDAENETASILMSGFRQWIMHNTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQK 422
Qy 421 ALEAALGKAGQOQGIILNALGQIASAAV 447
Db 423 ALEAALGKAGQOQGIILNALGQIASAAV 449

RESULT 9

US-09-564-479-4

; Sequence 4, Application US/09564479

; Publication No. US20030095973A1

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: WOMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: 032931/0230

; CURRENT APPLICATION NUMBER: US/09/564,479

; PRIOR FILING DATE: 2000-05-03

; PRIOR FILING DATE: 1999-05-03

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 583

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-564-479-4

Query Match 84.7%; Score 1866; DB 10; Length 583;

Best Local Similarity 100.0%; Pred. No. 3e-134;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 MSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTFFDDYKTAQATAYDTTFTSLSA 128
Db 1 MSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTFFDDYKTAQATAYDTTFTSLSA 60

Qy 129 DQAAVSLQDAVNTIKDTATDEETAIAEWETKNADAVKVGQAQITELAKYASDNOAIL 188
Db 61 DQAAVSLQDAVNTIKDTATDEETAIAEWETKNADAVKVGQAQITELAKYASDNOAIL 120

Qy 189 DSIQKLTSLDLOALLOQSVANNKKAELLKEMQDNPVVPKTPAIAGSLVDQDTATQ 248

Db 121 DSIQKLTSLDLOALLOQSVANNKKAELLKEMQDNPVVPKTPAIAGSLVDQDTATQ 180

Qy 249 IEKDGNAIRDAYFAGQNASGAVENAKNNSISNIDSAAATATAKTQIAEAKKFPDPSI 308

Db 181 IEKDGNAIRDAYFAGQNASGAVENAKNNSISNIDSAAATATAKTQIAEAKKFPDPSI 240

Qy 309 LQEAQWVIAEQKDLNKKIPADGSDVNPFTTGGSKQGGSSIGTSIRVSMILDDAENETA 368

Db 241 LQEAQWVIAEQKDLNKKIPADGSDVNPFTTGGSKQGGSSIGTSIRVSMILDDAENETA 300

Qy 369 SILMSGFRQWIMHNTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQKALEALGK 428

Db 301 SILMSGFRQWIMHNTENPDSQAQOELAAQAAKAAAGDSDSAAALADAQKALEALGK 360

Qy 429 AGQOQGIILNALGQIASAAVVSAGV 452

Db 361 AGQOQGIILNALGQIASAAVVSAGV 384

RESULT 10

US-09-841-132-436

; Sequence 436, Application US/09841132

; Patent No. US20020061848A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Probst, Peter

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.469C8

; CURRENT APPLICATION NUMBER: US/09/841,132

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599

; SOFTWARE: FastSeq for Windows Version 3.0/4.0

; SEQ ID NO 436

; LENGTH: 647

; TYPE: PRT

; ORGANISM: Chlamydia trachomatis serovar D

US-09-841-132-436

Query Match 20.9%; Score 461.5; DB 9; Length 647;

Best Local Similarity 29.7%; Pred. No. 1e-26;

Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;

Qy 2 VNPICGPIDETERTPPADLSAQGLEASAANKSAEQRIAGAEAKPKESKTDVSVERSL 61
Db 15 MNPIINGQI-----ANSETKESKSEA-----SPSASSSVSSWSFL 52

Qy 62 RSANVNLMSLADKLGIASSNSSSTSRSDVDSTTATPTPPPTFFDDYKTAQATAYDT 120
Db 53 SSAKHALISLRD--AILNKNSSPTDSLQLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110

Qy 121 IFTSTSLADIQAALVSLQDAVNTIKDTATDEETAIAEWETKNADAVKVGQAQITELAK 179
Db 111 LENATTLAEYETKMDIIMAAQDMERLAKQAEVTRIKEALQEKQ-----EVIDKLNOLVK 166

Qy 180 YASDNOAILDSLGLKLTSLDLOALLOQSVANNKKAELLKEMQDNPVVPKTPAIAGSLV 239
Db 167 LEKQNTKLTETLTSTDSADOIPAINSOLEINKNSADOIIXOLEGQNI---SYRAVLNAG 223

Qy 240 DOTDATATQIEKDGNAIRDAYFAGQNASGAVENAKNNSISNIDSAAATATAKTQIAEA 299
Db 224 EVIKASSEAGIKLQALQSVLDAGDQSQAAVLQAOQNNSPDNIAATKCLIDAAETKVNEL 283

Qy 300 QKK---FPDPSILQEAQWVIAEQKDLNKKIPADGSDVNPFTTGGSKQGGSSIGSI-- 354
Db 284 KQEHGTGTDSPLYKKAEEQISQAKDQIEIKPS-GSDIPIVGPS-GSAAGSAGSAGALKS 341

Qy 355 -----RVSMLLDDAENETASILMSGFRQWIMHNTENPDSQAQOELAAQAAKAAK---AA 406
Db 342 SNNSGRISLILDDVDNEMAIAMQGFMSLEQFVNNVPATAKELQAMEAQITAMSDQLVG 401

Qy 407 GDSAAALADAQKALEALGKAGQOQGIILNALGQIASAAVVSAG 451
Db 402 ADGELPAEIOAIKDALAQAL-KQPSDGLATAMGQVAFAAAKVGG 445

RESULT 11

US-09-841-260-139

; Sequence 139, Application US/09841260

; Publication No. US20030175700A1

; GENERAL INFORMATION:

; APPLICANT: Bhatia, Ajay

; APPLICANT: Probst, Peter

; APPLICANT: Stromberg, Erika Jean

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS

; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION

; FILE REFERENCE: 210121.515

; CURRENT APPLICATION NUMBER: US/09/841,260

; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 140

; SEQ ID NO 139

; LENGTH: 660

; TYPE: PRT

; ORGANISM: Chlamydia trachomatis

US-09-841-260-139

Query Match 20.6%; Score 453; DB 10; Length 660;
Best Local Similarity 28.9%; Pred. No. 4.6e-26;
Matches 136; Conservative 94; Mismatches 193; Indels 48; Gaps 14;

QY 6 GPGPIDETERTPPADLSAAGLEASAANKSABEARIAGAEAKPKESKTDVSERWSILRSV 65
DB 11 GPESVSSNQSNMPPINGQIASNETKSTKA-----SEASP--SASSSVSSWSFLSSAK 63

QY 66 NALMSLADKLGIIASSNSSSTSRK-ADVDTTATAPTPTTPTTDDYKTAQATYDTIFTS 124
DB 64 NALISLRD--AALNKNSPTDLSQLEASTSISTVRVAAKYDEAKSNFTAKSGLENA 121

QY 125 TSLADIQAALVSLQDVTNIKDTAATD-----EET-----ATAEWETKNADAVKVGQA 173
DB 122 KTLAEYETKMADLMAALQDWERLANGSDPSNNHTEEVNNIKKALEAQKDTID-----K 173

QY 174 ITELYASDNQAALDLSGLKTSFDLLQAALLQSVANNKKAABELLKEMQDNPVPGKTPA 233
DB 174 LNKLVTLQNKNSLTVLKTTSADQIPAINSQLINKNSADQIIKDLERQNI---SYEA 230

QY 234 IAQSLVDQTDATATQIEKQGNARDAVFAGQNASGAVENAKSNNSISNIDSAKAAIATAK 293
DB 231 VLTNAGEVTKASSEAGIKLQALQSIVDAGDQSOAAVLQAQNNSPDNIAATKELLIDAE 290

QY 294 TQIAEAQKK---FPDPSPILOAEQMVIOAEKDLKNIKPADGSDVNPNGTTVGGSKQOQSS 350
DB 291 TKVNELKQHTGLTDSPLVKAEQISQAQKDIQIKPS--GSDPIVGPFS--GSAASAGSA 348

QY 351 IGS1-----RVSMILDDAENETASILMSGFROMIHMENTENPDQSAQOELAAQARAA 403
DB 349 AGALKSNNSGRISLLDDVDNEMAALQGFRRSMIEQFNVNPNPAKELQAMEAQLTAM 408

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQOQILNALQIASAAVVSAG 451
DB 409 SDQLVGADGELPAEIOAIKDALLAQAAL--KOPSADGLATANGQVAFPAAKVGG 458

RESULT 12
US-10-007-693-139
; Sequence 139, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121-515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 139
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-007-693-139

Query Match 20.6%; Score 453; DB 13; Length 660;
Best Local Similarity 28.9%; Pred. No. 4.6e-26;
Matches 136; Conservative 94; Mismatches 193; Indels 48; Gaps 14;

QY 6 GPGPIDETERTPPADLSAAGLEASAANKSABEARIAGAEAKPKESKTDVSERWSILRSV 65
DB 11 GPESVSSNQSNMPPINGQIASNETKSTKA-----SEASP--SASSSVSSWSFLSSAK 63

QY 66 NALMSLADKLGIIASSNSSSTSRK-ADVDTTATAPTPTTPTTDDYKTAQATYDTIFTS 124
DB 64 NALISLRD--AALNKNSPTDLSQLEASTSISTVRVAAKYDEAKSNFTAKSGLENA 121

QY 125 TSLADIQAALVSLQDVTNIKDTAATD-----EET-----ATAEWETKNADAVKVGQA 173
DB 122 KTLAEYETKMADLMAALQDWERLANGSDPSNNHTEEVNNIKKALEAQKDTID-----K 173

QY 174 ITELYASDNQAALDLSGLKTSFDLLQAALLQSVANNKKAABELLKEMQDNPVPGKTPA 233
DB 174 LNKLVTLQNKNSLTVLKTTSADQIPAINSQLINKNSADQIIKDLERQNI---SYEA 230

QY 234 IAQSLVDQTDATATQIEKQGNARDAVFAGQNASGAVENAKSNNSISNIDSAKAAIATAK 293
DB 231 VLTNAGEVTKASSEAGIKLQALQSIVDAGDQSOAAVLQAQNNSPDNIAATKELLIDAE 290

QY 294 TQIAEAQKK---FPDPSPILOAEQMVIOAEKDLKNIKPADGSDVNPNGTTVGGSKQOQSS 350
DB 291 TKVNELKQHTGLTDSPLVKAEQISQAQKDIQIKPS--GSDPIVGPFS--GSAASAGSA 348

QY 351 IGS1-----RVSMILDDAENETASILMSGFROMIHMENTENPDQSAQOELAAQARAA 403
DB 349 AGALKSNNSGRISLLDDVDNEMAALQGFRRSMIEQFNVNPNPAKELQAMEAQLTAM 408

QY 404 K---AAGDPSAAALADAQKALEAALGKAGQOQOQILNALQIASAAVVSAG 451
DB 409 SDQLVGADGELPAEIOAIKDALLAQAAL--KOPSADGLATANGQVAFPAAKVGG 458

RESULT 13
US-09-841-132-299
; Sequence 299, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-299

Query Match 10.5%; Score 231.5; DB 9; Length 361;
Best Local Similarity 35.4%; Pred. No. 1.8e-09;
Matches 57; Conservative 30; Mismatches 61; Indels 13; Gaps 5;

QY 301 KKFDPSPILQEAQMVIOAEKDLKNIKPADGSDVNPNGTTVGGSKQOQSSIGSI----- 354
DB 2 QEIADSPLVKKAEEQINQAQDIQITIFS--GLDPIVGPFS--GSAASAGSAAGALKSSNNS 59

QY 355 -RVSMILDDAENETASILMSGFROMIHMENTENPDQSAQOELAAQARAAK---NAGDDS 410
DB 60 GRISLLDDVDNEMAALQGFRRSMIEQFNVNPNPAKELQAMEAQLTAMSDQLVGADGE 119

QY 411 AAAALADAQKALEAALGKAGQOQOQILNALQIASAAVVSAG 451
DB 120 LPAEIOAIKDALLAQAAL--KOPSADGLATANGQVAFPAAKVGG 159

RESULT 14
US-09-971-536-69
; Sequence 69, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Bloksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:39:25 ; Search time 8.31079 Seconds
(without alignments)
2831.945 Million cell updates/sec

Title: US-10-608-559-6

Perfection score: 2204

Sequence: 1 MVNPIGPGIDERTTPAD.....QGILNALQIASAAVVSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	184	8.3	1120	1	STFR_ECOLI
2	176.5	8.0	545	1	HTR5_HALN1
3	160	7.3	535	1	HTR1_HALN1
4	156	7.1	810	1	HTR4_HALN1
5	156	7.1	810	1	HTR4_HALN1
6	154.5	7.0	1238	1	SRCC_RHOCA
7	151	6.9	641	1	HRPK_PSESY
8	149	6.8	778	1	HTR6_HALSA
9	148.5	6.7	763	1	HTR2_HALN1
10	148.5	6.7	764	1	HTR2_HALSA
11	147.5	6.7	1528	1	SPAA_STRDO
12	147	6.7	505	1	FLJB_SALTY
13	146	6.6	778	1	HTR6_HALN1
14	145	6.6	1565	1	PAC_STRMU
15	145	6.6	1714	1	SVEP_DROME
16	144	6.5	881	1	PRY3_YEAST
17	144	6.5	1562	1	SPAP_STRMU
18	143.5	6.5	758	1	INMT_HUMAN
19	141.5	6.4	2542	1	TLN2_HUMAN
20	141	6.4	535	1	HTR1_HALSA
21	141	6.4	819	1	NRFC_HAEIN
22	139.5	6.3	774	1	STFL_LAMED
23	139	6.3	705	1	CWBA_BACSU
24	139	6.3	1637	1	MRSP_STRAU
25	138.5	6.3	1969	1	MYSA_CAEEL
26	138	6.3	1609	1	LMG1_HUMAN
27	136	6.2	1607	1	LMG1_MOUSE
28	135.5	6.1	1061	1	TRC4_ECOLI
29	135	6.1	1306	1	MSB2_YEAST
30	134	6.1	376	1	FLAC_VIBAN
31	134	6.1	490	1	CE05_ECOLI
32	134	6.1	564	1	M12_STRPY
33	134	6.1	8545	1	ANCI_CAEEL

RESULT 1

```

STFR_ECOLI
ID STFR_ECOLI STANDARD; PRT; 1120 AA.
AC P76072; P77560;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Side tail fiber protein homolog from lambdaoid prophage Rac.
GN STFR OR H1372.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97425617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
SEQUENCE FROM N.A.
RN STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -!- SIMILARITY: Belongs to the tail fiber family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000234; AAC74454.1; ALT_INIT.
CC EMBL; D90774; BAA14966.1; -.
CC EMBL; D90775; BAA14975.1; -.
CC PIR; G64887; G64887.
CC EcoGene; EG13370; stfr.
CC InterPro; IPR008969; CarboxypepD reg.
CC InterPro; IPR005003; Phage fiber.
CC InterPro; IPR005068; Phage fiber_2.
CC Pfam; PF03335; Phage_fiber_6.
CC Pfam; PF03406; Phage_fiber_2; 1.

```

```

34 133.5 6.1 1013 1 A60D_DROME
35 133.5 6.1 1411 1 TCOF_HUMAN
36 133 6.0 500 1 FLJB_SALAE
37 133 6.0 797 1 VGLX_HSVBB
38 133 6.0 1205 1 SMC_MYCTU
39 132.5 6.0 2492 1 TALA_DICDI
40 132.5 6.0 2541 1 TLN1_MOUSE
41 132 6.0 475 1 MDTM_MOUSE
42 132 6.0 757 1 IMMT_MOUSE
43 132 6.0 2843 1 APC_HUMAN
44 132 6.0 2845 1 APC_MOUSE
45 131.5 6.0 573 1 YEB8_YEAST

```

ALIGNMENTS

```
KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
SQ SEQUENCE 1120 AA; 113779 MW; 542E59D71EE795B4 CRC64;

Query Match
Best Local Similarity 8.3%; Score 184; DB 1; Length 1120;
Matches 102; Conservative 69; Mismatches 201; Indels 78; Gaps 14;

QY 8 GPIDETERTPPA---DLQAQGLEAAGAAKSAQAQIAGAEAKPKESKTSVERWILRS 63
Dd 91 GAWTEDDARPEALRRLELM---VEEVARNASAVAQNTAAKKSASDASTISAREAAHAA 147
QY 64 ANVALMSLADKLGIASNSSSSTRSADVDSTATPTPPPTFFDYKQQAQYDTIT 123
Dd 148 AADSARAASTAGQAAAS-SQAASSAGTAGTASTKATASKAASAAESKSAATAGAAKT 206
QY 124 S--TSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKYA 181
Dd 207 SETNASASIQSATSASTATTKASE-AATSARDAASKAKKSSSE-----TNASSA 257
QY 182 SDNQAILDSIGKLTSPDLLQAALLQSVANNKAAELKEMQDNVPVFGKTPAIAQSLVDQ 241
Dd 258 SSAASATAAGN-----SAKAAKTSETVARSSETAAGQSAAGAAAGSKTAAASASASA 310
QY 242 TDATATQIEKQGNADAFAGQNAAGAVENAKSNNSISNIDSAKAAIATATQIAEAK 301
Dd 311 TSA-----GOASATATAAGKSA-----ESAASASTATTKAGEATE 346
QY 302 KFPDPSILOAEQWVIQAEKDLKNIPKAGSDVPNPGTTVGGSKQQGSSIGSIRVSMLLD 361
Dd 347 Q-----ASAASASAAKTSSETNAKASETSABESSKTAASASSASASASAK--D 398
QY 362 DAENETASILMSFROMIHFNTENPDQSAQOQLAAQA-----RAAKAAGDSDSAAAL 415
Dd 399 EATROASAASSATTAATKATEAAGSATAAQAQSKTAESAAATRAETAAKRAEDIAAVAL 458
QY 416 ADAQKALEALGAKAGQOQGLNALGQIASA 445
Dd 459 EDASTT-----KKGIV-----QLSSA 474

RESULT 2
HTR5 HALN1 STANDARD; PRT; 545 AA.
AC Q48318; Q9HP85;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Halobacterial transducer protein V.
GN HTR7 OR HTPV OR VNG1759G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OS Halobacterium salinarium.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091, 2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Ballig N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=H.salinarium; STRAIN=S9;
RX MEDLINE=96275896; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
```

```
RA Oesterhelt D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC !- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction.
CC !- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC !- SIMILARITY: Contains 1 HAMP domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE005080; AAG19985.1; -.
DR EMBL; X95589; CAAG4842.1; -.
DR PIR; E84327; E84327.
DR PIR; T46811; T46811.
DR HSSP; P02942; 1QU7.
DR
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_Chmtaxis.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PRO0260; CHENTRNSDUCR.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1. Complete proteome.
KW Transducer; Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT DOMAIN 164 217 HAMP.
FT DOMAIN 236 476 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 545 AA; 57070 MW; FDD870389C2F428B CRC64;

Query Match
Best Local Similarity 8.0%; Score 176.5; DB 1; Length 545;
Matches 98; Conservative 79; Mismatches 195; Indels 87; Gaps 17;

QY 12 ETERTPPADLSAQGLEAAGAAKSAQAQIAGAEAKPKESKTSVERWILRSVAVNALMSL 71
Dd 118 EEERA-----EAERAREKAEQKQAEAE---QTAEASAKODARERSAEIEQLAADLSQ 169
QY 72 ADKLGIASSNSSSTSRGADVSTTATPTPPPT-FDYKQQAQYDTI-----FTST 125
Dd 170 ATEVG-ATLEASDGDLTARVDAITDNEAIEAVATVNDMLTMTERTIDEIOGFSTNVT 228
QY 126 SLADIQAALVSLQDAVTNIKDT-----AATDBE---TAIAAEWETKNADAVKVGQAQITE 176
Dd 229 ASREATAGAKEIQDASQTVSESVQBIAGTDQREQLSEVABEMDSYSATVEEVAATAQS 288
QY 177 LAKYASDNCAILDSIGKLT-----SFDLLQAALLQSVANNKAAELKEMQDNVPVFGK 230
Dd 289 VADTAADTTDVA-TAGKQTAEDAIDAIDAVQETWQTTVAVNDAEDLTTEIDD----- 340
QY 231 TPAIAQSLVDQDATATQIEKGNADAFAGQNAAG-----AVENAKSNNSI 279
Dd 341 IAEILSDTAEQTNMLA--LNANIEARAGSGGSGNGDGFVAVDEVKELATESQSAKDI 398
QY 280 SN-ISAAXAATATKQTAIEAQAQKFPDPSPIIQAEQVMVIAEKDUNKTKPADGSDVPNPG 338
Dd 399 AELIEEVQSQTATVEEI-----RVAEQRVNDGAAAVEETVDAPGAVTENIQ 445
QY 339 TTVGSGKQGGSSIGSIRVSMLLDDEAENETASILMSFGFRQIMHFNTEPDSQAQOQLAA 398
Dd 446 EFTDGVQE-----ISQAMDEQAQRSERVSS-----VDDIATISQATAD 484
QY 399 QARAAKAGDSDSAAALADAAQKALEALGKAGQOQGLN 437
Dd 485 RAENTVSAASEEQ-TASITEVTSISQSLAAQAQADTLERLN 522
```



```

RESULT 3
HTR1 HALN1 STANDARD; PRT; 535 AA.
ID HTR1 HALN1
AC P33741; Q9HPP6;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sensory rhodopsin I transducer (HTR-I) (Methyl-accepting phototaxis
protein I) (MPP-I)
GN HTR1 OR HTR1 OR VNGI1659G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OS Halobacterium halobium.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
CX NCBI_TaxID=64091, 2242;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-10; 350-377 AND 457-476.
RC SPECIES=H.halobium; STRAIN=FX5R;
RX MEDLINE=93101637; PubMed=1465418;
RA Yao V.J., Spudich J.L.;
RA "Primary structure of an archaeobacterial transducer, a
RT methyl-accepting protein associated with sensory rhodopsin I.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:11915-11919(1992).
CC -!- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR
CC SENSORY RHODOPSIN I (SR-I) TO THE FLAGELLAR MOTOR. RESPONDS TO
CC LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE005075; AAG19913.1; -
CC EMBL; L05603; AAA72315.1; -
CC PIR; A47190; A47190.
CC PIR; E84318; E84318.
CC HSSP; P02942; 1QU7.
CC InterPro; IPR004089; Chmtaxis_transd.
CC InterPro; IPR003660; HAMP.
CC InterPro; IPR004090; Me_chmotaxis.
CC Pfam; PF00672; HAMP; 1.
CC Pfam; PF00015; MCPsignal; 1.
CC PRINTS; PR00260; CHEMTXNSDUCR.
CC SMART; SM00304; HAMP; 2.
CC SMART; SM00283; NA; 1.
CC PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
CC PROSITE; PS50885; HAMP; 2.
KW Transducer; Photoreceptor; Transmembrane; Methylation;
KW Complete proteome; Repeat.
FT INIT MET 0 0
FT DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT TRANSNM 14 28

```

```

FT DOMAIN 29 38 EXTRACELLULAR (POTENTIAL).
FT TRANSNM 39 54 POTENTIAL.
FT DOMAIN 55 535 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 54 106 HAMP 1.
FT DOMAIN 148 201 HAMP 2.
FT DOMAIN 220 458 METHYL-ACCEPTING TRANSDUCER.
FT MOD_RES 265 265 METHYLATION.
FT MOD_RES 272 272 METHYLATION.
FT MOD_RES 279 279 METHYLATION.
FT MOD_RES 463 463 METHYLATION.
FT MOD_RES 472 472 METHYLATION.
SQ SEQUENCE 535 AA; 56544 MW; B9945E4F66A9D091 CRC64;

Query Match 7.3%; Score 160; DB 1; Length 535;
Best Local Similarity 19.1%; Pred. No. 0.15;
Matches 90; Conservative 81; Mismatches 189; Indels 110; Gaps 15;

QY 8 GPIDETERTPPADLSAAGLEASAANKSAE--AQRAGAEAKPKE---SKTDSVERW--- 58
Db 103 GRUNEMERT-RADLEETQAEATAREAEAKQAEQAQAAREARELAATYQDTAKRYGETM 161
QY 59 -----SILRSVAVNALMSLADKLGIASSSSSST 86
Db 162 EAAATGDLTORVDVTDHEAMETVGTAFNOMDDLOATVTVTTVADEIEAKTERMSET- 220
QY 87 SRADYVDST---TATAPTPTPTFDYKTAQAYDTIFTSTSLADIQAALVSLQDAVTN 143
Db 221 --SADIASEAGDVFEAVKIESQANDQRTELDGAAD-----DVQQVSAEAEIAAT 269
QY 144 IKDTATATDEETATAAE-----WETKNADAVKVGAKITELAKYASDNOAIL 188
Db 270 IDLASRESDVATASDAARDSKSKALDEMSSITEVDDAVGQVEQLRDQVABITDVVI 329
QY 189 DSGKUTSFDLQALQLQSVANN-----KAEALLKEMQNVVPGTKPTAQAS 237
Db 330 TDIGEQTNLNALNASIEAARAGNADGDFSVVADEVKOLAETQDR---ANEIAAVVEK 386
QY 238 LVDQTDATATQIEKGNAINRDVAFQGNASGAVENAKSNNSISNIDSAKAAIATAKTQIA 297
Db 387 VTAQTDVNTASIQOTTRV-----EGSETVE-----STLRDIRTIADSIASVNSID 434
QY 298 EAQKKEPDSPILOEAEQWIAEKDKLNKIPADGSDVPNPGTTVGSGSKQGSSTIGSIRVS 357
Db 435 EIQTTSQAEATVQSTATSVERVAGLSDDDTALASDAE---SAVIGORESAEIAA--- 487
QY 358 MLDDAENETASIMSGFROMIMHFNENPENSQAQOELAAQARAKAAG 407
Db 488 -SLEQFQNTAVEQLQS-----RVASFVATEDSETAGSGVEQPVVRAGADG 532

RESULT 4
HTR4 HALN1 STANDARD; PRT; 810 AA.
ID HTR4 HALN1
AC Q9HPB4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Halobacterial transducer protein IV.
GN HTR5 OR HTPIV OR VNGI1760G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,

```

```

RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction (by similarity).
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005080; AAG19986.1; -.
DR PIR; F84327; F84327.
DR HSSP; P02942; 1QU7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS00111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 2.
DR Transducer; Transmembrane; Complete proteome; Repeat.
KW DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 POTENTIAL.
FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 344 POTENTIAL.
FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 345 397 HAMP 1.
FT DOMAIN 439 493 HAMP 2.
FT DOMAIN 512 748 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 810 AA; 85219 MW; 4BF36E4B7D2BD80 CRC64;

Query Match
Best Local Similarity 7.1%; Score 156; DB 1; Length 810;
Matches 102; Conservative 73; Mismatches 170; Indels 118; Gaps 22;

QY 27 EASAANKSA-EAQRIGAEAKPKESKTSVERWSILRSVAVNLMADKLGIASSNSSSS 85
Db 398 EAAAREQATEAQDAEAE---RERAEADAREADAKADAEL-----AAELEAQ 444

QY 86 TERSADVSTTATAPPPPTFDYKTAQAVDTFTSTSLADIQAALVSLQDAVTNIK 145
Db 445 AERYSDVMAACADGDLTRRMPADDTNEAAMAIAASF-NEMLAQWHEHTIIDIQE---FA 499

QY 146 DTAATDEETAIAAEWETKKNADAVKVGQIT---ELAKYASDNQAILDSL-GKLTSPDLL 200
Db 500 DAVATASEEA-----EVGAADAREASQVSEVQETAGADEQRNMLDVTSGEMT--DL- 551

QY 201 QAALQSIVANNKAAELKEMQNPVPGKTPAIAQSLVDQDTATATQIEKGNARIDAY 260
Db 552 -SAAIEVA-----ASADSVASHSHQTA-ETIARDGE-----580

QY 261 FAGQNSGAVENAKSNNSISNIDSAKAAIATATQTAABAKKFPD--SPILQAEQWVIQ 318
Db 581 ---QTAEADAEI-E-RSLTVQEAIDATVQNVBALDDQMAEI-SEIVDLISDIAEQTNMLAN 634

QY 319 AE-----KDLKNIKPADGSDVENPQTTVGGKQGGSSIGSIRVSM 358
Db 635 ANTEARADKSGDGFVAIVDEVKDLAEETQESAGDIERRITEV--QSQTATVAERAAE 692

QY 359 LLDDAENETASILMSGRFMI-HMFNT-----ENPDSQAQOE-----395
Db 693 ESMADGIDAVEEVVDATAFTAVSDHAETDTGVQETISDITDDQAASTEBAVSMTEVADLSD 752

QY 396 -LAAQARAANKAGDSSAAALADAKALEALGKAGQCGQILN 437
Db 753 STAGEAQSVAAABEQ-AASMSEISDSVESLSGQAEQLKALLS 794

```

```

RESULT 5
HTR4_HALSA
ID HTR4_HALSA STANDARD; PRT; 810 AA.
AC Q48317;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Halobacterial transducer protein IV.
GN HTR5 OR HTRIV.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_taxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S9;
RX MEDLINE=96275996; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
RA Oesterhelt D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction.
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95589; CAA64841.1; -.
DR PIR; T46810; T46810.
DR HSSP; P02942; 1QU7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS00111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 2.
DR Transducer; Transmembrane; Repeat.
KW DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 POTENTIAL.
FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 324 344 POTENTIAL.
FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 345 397 HAMP 1.
FT DOMAIN 439 493 HAMP 2.
FT DOMAIN 512 748 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 810 AA; 85207 MW; 37B0F6D46A3D9BA CRC64;

Query Match
Best Local Similarity 7.1%; Score 156; DB 1; Length 810;
Matches 101; Conservative 74; Mismatches 170; Indels 118; Gaps 22;

QY 27 EASAANKSA-EAQRIGAEAKPKESKTSVERWSILRSVAVNLMADKLGIASSNSSSS 85
Db 398 EAAAREQATEAQDAEAE---RERAEADAREADAKADAEL-----AAELEAQ 444

QY 86 TERSADVSTTATAPPPPTFDYKTAQAVDTFTSTSLADIQAALVSLQDAVTNIK 145
Db 445 AERYSDVMAACADGDLTRRMPADDTNEAAMAIAASF-NEMLAQWHEHTIIDIQE---FA 499

QY 146 DTAATDEETAIAAEWETKKNADAVKVGQIT---ELAKYASDNQAILDSL-GKLTSPDLL 200
Db 500 DAVATASEEA-----EVGAADAREASQVSEVQETAGADEQRNMLDVTSGEMT--DL- 551

```



```

RESULT 9
ID HTR2 HALNI STANDARD; PRT; 763 AA.
AC Q9HP81;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
DE protein II) (MPP-II).
GN HTR2 OR VNG1765G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welte R., Goo Y.A.,
RA Lethausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Eberhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,
RA "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Transduces signals from the phototaxis receptor sensory
CC rhodopsin II (SR-II) to the flagellar motor. Responds to light
CC changes through the variation of the level of methylation. Also
CC acts as a chemotransducer (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE005080; AAC19989.1; -.
CC F01; A84328; A84328.
CC HSP; F02942; 1Q07.
CC InterPro; IPR004089; Chmtaxis_transd.
CC InterPro; IPR003660; HAMP.
CC Pfam; PF00672; HAMP; 1.
CC Pfam; PF00015; MCFsignal; 1.
CC SMART; SM00304; HAMP; 1.
CC SMART; SM00283; MA; 1.
CC PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
CC PROSITE; PS00885; HAMP; 2.
CC Transducer; Photoreceptor; Chemotaxis; Transmembrane; Methylation;
CC Complete proteome; Repeat.
CC INIT_MET 0 0 BY SIMILARITY.
CC DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 16 36 POTENTIAL.
CC DOMAIN 37 277 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 278 297 POTENTIAL.
CC DOMAIN 298 763 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 302 354 HAMP 1.
CC DOMAIN 396 449 HAMP 2.
CC TRANSMEM 468 704 METHYL-ACCEPTING TRANSDUCER.
CC PROSITE 763 AA; 78911 NW; CF7A8FF04DF309A CRC64;
SQ SEQUENCE 763 AA; 6.7%; Score 148.5; DB 1; Length 763;
Query Match 6.7%; Score 148.5; DB 1; Length 763;
Best Local Similarity 22.3%; Pred. No. 0.93;
Matches 105; Conservative 72; Mismatches 190; Indels 103; Gaps 20;
8 GPIDETPTPADLSAQGLEASANKSAB-----AQRAGAEAKPK-----SKTDSVERV 58

```

```

Db 336 GTLAESFRSMRDSLSLTDARATARAEDAREDAEQRAAEAREDAEAARKDAQETA 395
Qy 59 SILRSAV-----NALMSLADKLGIIASSNSSSSSTSSA-----DVSSTTATAPT 101
Db 396 RALESAAADYEALTAVAD--GDLTRRVDAQRDHDAVARIGHALNDMLDDIETSVAAA-- 451
Qy 102 PPPFTFDYKQA-----QTAAYDTFTSTSLADTQAALVSLQDAVTNIKOTAADEETAI 156
Db 452 ---TAFSDHVSDDAAQRAVEADAGDAIDAGT---DVSTAVDEISDGATE-----QTDRLHEV 500
Qy 157 AAWEETKNADAVKYGQITELAKYASDNQATLDSLSGLKLTSPDLLQALLOQVANNKAAE 216
Db 501 AGEVDDLSASAEVAETVASLADTAGQAASAVDD--GRQATEDAVET--MDVDADAEAAA 557
Qy 217 LKEMQDNPPVP--GKTPAIAQSLVDQTDATQIEKGNIRDAIRDAYFAGQNASGAVENA-- 273
Db 558 DAMDALDSEMDIGEIVDVIAADIADQINMLAL-----NASIEAARTGADGDGFVVADE 611
Qy 274 -----KSNNSISNIDSAAKAIATAKTQIABAQKKFPDPSILOEABQWVIOAEKDLKNIK 327
Db 612 VKTLAESRDAEDIESRLALQGVSDVADMEATSDT--VSDGRATVGDATALDDV-- 668
Qy 328 PADGSDVNPCTTVGGSKQGGSSIGSRVSMLLDDAENETASIIILMSGFROMIHMENTNP 387
Db 669 -----VSFVADITDAAGEIRAA---TDRQHAAASRVASAVDEV----- 703
Qy 388 DSAQAQELAAQARA--AKAAGD-----DSAAAALADAQKALEAALGK 428
Db 704 --AGISQETAQAQATAVADSAAATQDTLSSVDDAAADLADRAAALDDLLAE 751

RESULT 10
ID HTR2 HALSA STANDARD; PRT; 764 AA.
AC P71410;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
DE protein II) (MPP-II).
GN HTR2 OR HTRII.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FLX15;
RX MEDLINE=96323203; PubMed=8710852;
RA Zhang W., Brooun A., Mueller M.M., Alam M.;
RT "The primary structures of the Archaeon Halobacterium salinarum blue
RT light receptor sensory rhodopsin II and its transducer, a methyl-
RT accepting protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:8230-8235(1996).
RN [2]
RP FUNCTION.
RX PubMed=9515936;
RA Hou S., Brooun A., Yu H.S., Freitas T., Alam M.;
RT "Sensory rhodopsin II transducer HtrII is also responsible for serine
RT chemotaxis in the archaeon Halobacterium salinarum.";
EL J. Bacteriol. 180:1600-1602(1998).
CC -!- FUNCTION: Transduces signals from the phototaxis receptor sensory
CC rhodopsin II (SR-II) to the flagellar motor. Responds to light
CC changes through the variation of the level of methylation. Also
CC acts as a chemotransducer.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


Db 188 KKAQYEQDLAANKAEVSLMRKPRPIYAK---LAQNKDL--RAIQANSDSQAY 242
 Qy 255 -AIRDAY-----FAGQNASGAVENAKSNNSISNIDSAXAIATATKTQIAEAKKFPD 305
 Db 243 AAAREAYDKWARVQAANAANKAYEEALAAANTAKN-DQIKAEIEAQORSAKA----- 295
 Qy 306 SPILQEAQOMVIOAEKDKLNKIPADGSDVPNPGTIVGSKQOGSSISIRVSMLLDAEN 365
 Db 296 -----DYEAKLAQYEQDKLAAQAAGNAANEADYQAKKAAYEQELARV-----QAAN 340
 Qy 366 ETASILMSGFQMIHMENTEN-----PQSAQAQBELAQARAKAAGDDSAALADAKA 421
 Db 341 AAAR--QAYEQALANSKAKNAQIATNEALQON--AQAKA-----DYEAKLAQYQKO 388
 Qy 422 LEAA-LKQAGQOQQILNALG-----QTASAUVSA 450
 Db 389 LAAQAQSGNAANEADYQEKLAAYEKELARVQAANAQA 427

RESULT 12

FLJB_SALTY STANDARD; PRT; 505 AA.
 ID FLJB_SALTY STANDARD; PRT; 505 AA.
 AC P52616; P97159;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phase-2 flagellin.
 GN FLJB OR H2 OR STM2771.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SL 375;
 RA MEDLINE=95325331; PubMed=7541401;
 RX Vaneas R.A., Joys T.M.;
 RT "Molecular analyses of the phase-2 antigen complex 1,2,. . of
 Salmonella spp.";
 RL J. Bacteriol. 177:3863-3864 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2.";
 RL Nature 413:852-856 (2001).
 RN [3]
 RP SEQUENCE OF 1-37 FROM N.A.
 RX MEDLINE=82049491; PubMed=6271461;
 RA Silverman M., Zieg J., Mandel G., Simon M.;
 RT "Analysis of the functional components of the phase variation
 system.";
 RL Cold Spring Harb. Symp. Quant. Biol. 45:17-26 (1981).
 RN [4]
 RP SEQUENCE OF 482-505 FROM N.A.
 RC STRAIN=SUJ2353;
 RA Mingorance J., Tanaka S., Tomimaga A., Enomoto M.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Flagellin is the subunit protein which polymerizes to
 form the filaments of bacterial flagella.
 CC -!- MISCELLANEOUS: Individual salmonella serotypes usually alternate
 between the production of 2 antigenic forms of flagella, termed
 Phase-1 and Phase-2, each specified by separate structural genes.
 CC -!- SIMILARITY: Belongs to the bacterial flagellin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U17177; AAC43354.1; --
 DR EMBL; AE008826; AAL21657.1; --
 DR EMBL; V01370; CAA24655.1; --
 DR EMBL; D26168; BAA05156.1; --
 DR ScvGene; SG10564; fljB
 DR InterPro; IPR001029; Flagellin C.
 DR InterPro; IPR001492; Flagellin N.
 DR Pfam; PF00700; Flagellin C; 1.
 DR Pfam; PF00669; Flagellin N; 1.
 DR PRINTS; PR00207; FLAGELLIN.
 DR ProDom; PD000316; Flagellin C; 2.
 KW Flagellum; Complete proteome.
 FT INIT_MET 0 BY SIMILARITY.
 FT CONFLICT 37 I -> S (IN REF. 3).
 SQ SEQUENCE 505 AA; 52404 MW; 3A5CC404AF7AF88B CRC64;
 Query Match 6.7%; Score 147; DB 1; Length 505;
 Best Local Similarity 20.8%; Pred. No. 0.68;
 Matches 92; Conservative 62; Mismatches 166; Indels 122; Gaps 19;
 Qy 23 AQLLEASANKSAEPAQRIAG-----AEAPKESKTDVSVERWSILRSAYNALMSLADKL 75
 Db 32 SSGLRINSKDDAAGQAIAANRFTANIKGLTQASRNANDG:SIQTTTEGALNEINNLRV 91
 Qy 76 GIASSSSSTSRSAVDSTATAP-----TPPPTFDYKTAQATYDIFTST- 126
 Db 92 RELAVOSANSTNSQSDLDLSQAEITQRINIDRVSGQTQFNGVKVLAQDNTLTQVGAND 151
 Qy 127 --LADIQALVSLQ---DAVT---NIKDTAAT---DEETATAE 159
 Db 152 GETIDIDLKQINSQTLGLDSLNVQKVDVKTAVTTKAYANNGTTLDSGLDDAAIKAT 211
 Qy 160 WETKNADAVKVGQITELAKYASDNQAILDSLGLKLTSTFD-----LLQ 201
 Db 212 GGTNGTASVTGGA-----VKFDADNNKYFVTIGFTGADAAKNGDYEVNVATDGTVTLAA 266
 Qy 202 AALLQSVANNKKAELLKEMQDNFV---PGKTPAQSLVDQDTATATOI-----EKD 252
 Db 267 GATKTNFAGATTKTEVQELKDTFAVVSADAKNALIAGG-VDATDANGAELVKMSYTDKN 325
 Qy 253 GNAR-----DAYFAG--QNASGAVENAKSN-----NSISNIDSAAKAI 289
 Db 326 GKTIEGGYALKAGDKYAADYDEATGAIK-AKTSYTAADGTTKTAANQLGGVD-GKTEV 383
 Qy 290 AT-----AKTQIAEAKKFPDPSPIQEAQOMVIOAE---KOLKNIKPA 329
 Db 384 VTIDGKTYNASKAGHDFKAQPELAEEAAKTENP-LQKIDAALAOVDLRLSDLGAVQNR 442
 Qy 330 DGSDDVNEGTGSGKQOGSSI 351
 Db 443 FNSAITNLGNTVNNLSEARSRI 464
 RESULT 13
 HTR6_HALN1
 ID HTR6_HALN1 STANDARD; PRT; 778 AA.
 AC Q9HR92;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Halobacterial transducer protein VI.
 GN HTR4 OR HTPVI OR VNG806G.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteriae; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Dawson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt T., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction (by similarity).
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005022; AAG19266.1; -.
DR EMBL; F84237; F84237.
DR HSSP; F02942; IQU7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR004090; Me_chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDCUR.
DR SMART; SM00304; HAMP; 2.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS01111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 2.
DR Transducer; Transmembrane; Complete proteome; Repeat.
FT DOMAIN 1 26 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 27 47 POTENTIAL.
FT DOMAIN 48 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 317 POTENTIAL.
FT DOMAIN 318 778 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 318 370 HAMP 1.
FT DOMAIN 412 465 HAMP 2.
FT DOMAIN 484 720 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 778 AA; 82076 MW; D7522787274FA79 CRC64;

Query Match 6.6%; Score 146; DB 1; Length 778;
Best Local Similarity 23.5%; Pred. No. 1.3;
Matches 103; Conservative 63; Mismatches 166; Indels 106; Gaps 18;

QY 20 DLSAQGLEAANKSAQAQIAAGAKPKSKTD-----SVERWSILRSA----- 64
Db 378 DAFAARSEAAQAQADAAQAQ-ACFAARERSEQAARLETTAFSEFMTVAAGDLTVR 436

QY 65 VNALMSLADKLGTASSNSSSSSTSRSDVSDTATPTPTPTDXYTQATYDIFTS 124
Db 437 LDADVQAAMADIAAAYNEWAADMEATIDVAF-----DEVATASTDASD----- 483

QY 125 TSLADIQAALVSLQAVNTIKDTAA--TDETAIAAEWETKNADAVKVGQITELAKYAS 182
Db 484 -SAAVQGTGRDVSADVGRIDRADQDQDLQEAIVASETDEMSATIEEVAASAOVAB-TS 541

QY 183 DNQAILDSGLKLTDFLLQALQSVANNKAAELLKEMQNPVPGKTPALAQSLVDQT 242
Db 542 QRAAALGDG-----QAAQDAVAQ-----LEEIED-----ETQA-AATAVDL 579

QY 243 DATATQTEKGNARD-----AYFAGQNASGAVENAKSNNSINISDAKA 287
Db 580 EAKVSEIETIVAAITDIAEQTNMLNANIEAARADQGDGFVAVDEVKDLA--DESKA 637

```

```

QY 288 ATATAKTIQAERAKKFPDPILOEABQWIOAEKDLKNIKPADGSDVNPFTVGGSKQ 347
Db 638 SAAETALVAEVR-----AQETSVAAMDRIQERVSDGVE-----TVSETERS 680
QY 348 GSSIGSIRVSMELDDAENETASILMSGFRQMTHTENTENPDQAAQCOELAAQARAAGA 407
Db 681 LSEIAG-RI-----AEADTG-----VQETSNAMDDQAAASVSDVTTAVGDAALG 723
QY 408 DSAALAAJADAQKALEAA 425
Db 724 EETATEAESTADAAAEQA 741

RESULT 14
PAC STRMU STANDARD; PRT; 1565 AA.
AC P11657;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pac protein precursor.
GN PAC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]_TaxID=1309;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-48.
RC STRAIN=MT8148 / Serotype C;
RX MEDLINE=89343654; PubMed=2761390;
RA Okahashi N., Sasakawa C., Yoshikawa M., Hamada S., Koga T.;
RT "Molecular characterization of a surface protein antigen gene from
RT serotype c Streptococcus mutans, implicated in dental caries."
RL Mol. Microbiol. 3:673-678(1989).
CC -!- FUNCTION: SURFACE PROTEIN ANTIGEN IMPLICATED IN DENTAL CARIES.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -!- SIMILARITY: STRONG TO AGGLUTININ RECEPTOR OF S.SANGUIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14490; CAA32652.1; -.
DR PIR; S04729; S04729.
DR PDB; 1JWW; 17-JUL-02.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR TIGRfam; TIGR01167; LPXTG_anchor; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Antigen; Cell wall; Peptidoglycan-anchor; Repeat; Signal;
KW Dental caries; 3D-structure.
FT SIGNAL 1 38
FT CHAIN 39 1535 PAC PROTEIN.
FT PROPEP 1536 1565 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 219 464 3 X TANDEN REPEATS, ALA-RICH.
FT SITE 1532 1536 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1535 1535 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 1565 AA; 170781 MW; 4C3B05C809D0C32A CRC64;

Query Match 6.6%; Score 145; DB 1; Length 1565;
Best Local Similarity 23.6%; Pred. No. 3.4;
Matches 107; Conservative 71; Mismatches 202; Indels 74; Gaps 22;

QY 11 DEFETPPADLSAQGLE-ASAANKSAEAOIAGAAKPKSKTDSVERWSILRSVNAIM 69
Db 39 DETTTSDVDTKVVGQTGNPATNIPDAQGSASKEAQSKTLE-----RQWHTIE 90
QY 70 SLADKLGIASSNSSS-----STSRSDVDSTTATAP-----TPPPPTFDYKTAQYDITF 122

```


Db 91 VPKTDLQAAKDAKSGVNVQDADVNGTVPKTPBEAVQKETEIKEDYTKQAEIDIKT-- 148
 Qy 123 TSTSLADTQAAVLSTQDAVTNFK-DTAIDEB-----TAAAEWETKN-ADAVKVGQAQT 175
 Db 149 TDQYKSDAAH-----EAEVAKTKAKNQATKEQYKDMAHKAEEVERINAANAASKTAYEA 204
 Qy 176 ELAKVASDNOALDLSGLKLTSDLLQALLOSVAANNKAAELLKEMQDNPVPGKTPAIA 235
 Db 205 KLAQY-----QADLAAVQKTNAN--QAAVOKALAYQ--AEUKRVQEAN-----AAAKAY 252
 Qy 236 SLDVDTQDATAQIQEKDGNINRDAYFAGQNASGAVEN-----AKSNNSISNDSAKAA--- 288
 Db 253 DTAVAANNAKNTETAAANEETRK-----RNATAKAEYETKLAQYQAEIKRVQEAANAENA 307
 Qy 269 -----IATAKTIQIAEAKKFPDSPILQAEQVWVQAEKDLKNIK-PAGDSVDPN-----P 337
 Db 308 DYQAKLTAYQTELARVQKANADAKATYEA-----AVAAANNAKNAALTAENTAIKQENENAK 363
 Qy 338 GTTVGSGKQQSSIGSIRVSMLLDDAENETASILMSGFROMHMENTENPDSQAQQAQELA 397
 Db 364 ATYEAALQYEAADLAAVKKNANNAEADYQAK---LTAYQTELARVQKANADAKAYEA 420
 Qy 398 AQARAAKAA--GDSAAAL-ADAQKALEALGK 428
 Db 421 AANNAANAALTAENTAIAKKNADAKADYEAKLAK 454

RESULT 15

SYEP_DROME STANDARD; PRT: 1714 AA.
 AC P28668; Q3VCF5;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bifunctional aminoacyl-tRNA synthetase [includes: Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate-tRNA ligase); Prolyl-tRNA synthetase (EC 6.1.1.15) (Proline-tRNA ligase)].
 GN AATS-GLUPRO OR CG5394.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=92097547; PubMed=1756734;
 RX Cerini C., Kerjan P., Astier M., Gratecos D., Mirande M., Semeriva M.;
 RA "A component of the multisynthetase complex is a multifunctional aminoacyl-tRNA synthetase."
 RL EMBO J. 10:4267-4277(1991).
 [2]
 SEQUENCE FROM N.A.
 RP MEDLINE=97217441; PubMed=9063462;
 RX Cerini C., Semeriva M., Gratecos D.;
 RA "Evolution of the aminoacyl-tRNA synthetase family and the organization of the Drosophila glutamyl-prolyl-tRNA synthetase gene. Intron/exon structure of the gene, control of expression of the two mRNAs, selective advantage of the multi-enzyme complex."
 RL Eur. J. Biochem. 244:176-185(1997).
 [3]
 SEQUENCE FROM N.A.
 RP STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Godok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.I., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeqwan C., Jialali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laske P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP + diphosphate + L-glutamyl-tRNA(Glu).
 CC -!- CATALYTIC ACTIVITY: ATP + L-proline + tRNA(Pro) = AMP + diphosphate + L-prolyl-tRNA(Pro).
 CC -!- SUBUNIT: Component of the multisynthetase complex which is comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl, arginyl, and aspartyl-tRNA synthetases as well as three auxiliary proteins, p18, p48 and p43.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -!- SIMILARITY: Contains 6 WHEP-TRS domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M74104; AAA28594.1; --
 EMBL; U59923; AAC47469.1; --
 EMBL; AE003745; AAF56211.1; --
 EMBL; S18244; S18644.
 HSP; P00962; IGTR.
 DR FlyBase; FBgn005674; Aats-glupro.
 DR InterPro; IPR004526; GltX arch.
 DR InterPro; IPR000924; Glu tRNA-synt_1c.
 DR InterPro; IPR004046; GST-Cterm.
 DR InterPro; IPR004154; HGTF_anticonodon.
 DR InterPro; IPR004499; ProS_fam1.
 DR InterPro; IPR002314; tRNA-synt_2b.
 DR InterPro; IPR001412; tRNA-synt_1.
 DR InterPro; IPR002316; tRNA-synt_pro.
 DR InterPro; IPR006195; tRNA_ligase_II.
 DR InterPro; IPR000738; WHEP-TRS.
 DR Pfam; PF00043; GST_C; 1.
 DR Pfam; PF03129; HGTP_anticonodon; 1.
 DR Pfam; PF00749; tRNA-synt_1c; 1.
 DR Pfam; PF03950; tRNA-synt_1c_C; 1.

DR Pfam; PF00587; tRNA-synt_2b; 1.
DR Pfam; PF00458; WHEP-TRS; 6.
DR PRINTS; PR00987; TRNASYNTHGLU.
DR PRINTS; PR01046; TRNASYNTHPRO.
DR TIGRFAMS; TIGR00463; gltx arch; 1.
DR TIGRFAMS; TIGR00408; PROS_fam_I; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
DR PROSITE; PS00762; WHEP-TRS; 6.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Multifunctional enzyme; Repeat.
FT DOMAIN 170 754 GLUTAMYL-TRNA SYNTHETASE.
FT DOMAIN 755 800 WHEP-TRS 1.
FT DOMAIN 827 872 WHEP-TRS 2.
FT DOMAIN 901 946 WHEP-TRS 3.
FT DOMAIN 980 1025 WHEP-TRS 4.
FT DOMAIN 1055 1100 WHEP-TRS 5.
FT DOMAIN 1129 1173 WHEP-TRS 6.
FT DOMAIN 1174 1180 POLY-GLY.
FT DOMAIN 1207 1714 PROLYL-TRNA SYNTHETASE.
FT SITE 209 220 "HIGH" REGION.
FT SITE 438 442 "KMSKS" REGION.
FT BINDING 441 441 ATP (BY SIMILARITY).
FT CONFLICT 102 106 TSPLP -> DKSIA (IN REF. 3).
FT CONFLICT 233 234 VC -> AF (IN REF. 3).
FT CONFLICT 341 345 NTACA -> KYCVR (IN REF. 3).
FT CONFLICT 583 583 K -> R (IN REF. 3).
FT CONFLICT 692 692 L -> A (IN REF. 3).
FT CONFLICT 753 753 T -> S (IN REF. 3).
FT CONFLICT 802 802 T -> S (IN REF. 3).
FT CONFLICT 873 873 P -> T (IN REF. 3).
FT CONFLICT 887 887 G -> V (IN REF. 3).
FT CONFLICT 1201 1201 P -> PA (IN REF. 3).
FT CONFLICT 1461 1461 MISSING (IN REF. 3).
FT CONFLICT 1587 1587 G -> V (IN REF. 3).
SQ SEQUENCE 1714 AA; 189197 MW; 6PB8C58045E48A8C CRC64;

Query Match 6.6%; Score 145; DB 1; Length 1714;
Best Local Similarity 19.4%; Pred.No. 3.7;
Matches 99; Conservative 75; Mismatches 193; Indels 142; Gaps 19;

QY 7 PGPI-----DETERPPADLSAQGLEASAAKSAEQAQRIAG-----A 43
Db 701 PSFIVLPSIPDGHKOVPTSGLVKNAPDAKATKASPSVSSGQASQLDSQITQQGLVLR 760
QY 44 EAKPKSKTDSVERWSILRSANVALMSLADKLGIASNSSTSRSDVDSTTAT-APT 102
Db 761 DLKSKAAKQID-----VAVKKLLAL-----KADYKSATGKDWKPGQTSATSAPVP 807
QY 103 PPFTDDYKTAQACTAYDTITFTSLADIQAALYSLODAVTNIKDTAATDEETAIA----- 157
Db 808 -----AASSSANDAV-----SVNASIVKQGLVREDLKGKASKPEIDAAVKTL 852
QY 158 ---AEMETKNADAVKVGQAITEAKYASDNQAI--LDSLGKLTSDLLQALQLQSVANN 212
Db 853 ELKQYKTLTGQDWKGTVPPTAAPSASAPSVGNDSDVAQILS-----QITAGD 903
QY 213 KAAELLKEMODNPVVPFGKTPAIAQSLVDQTDATQIEKDGNAIRDAYFAGONASGAVEN 272
Db 904 KVRELKSAKADKATVDAAVKTLTLKADYKAAAT-----GSDWKPGTTAPAPAAPVKV 956
QY 273 AKSNN-----SISNIDSKAIAIAKQIAEAKKFPDPSPILOEQMVQAEKDLKNI-- 326
Db 957 KQKNPDPAVLVTNLLNKIAQGGDKIQRLKSAKSEKSLVEAEVKLLALKTDYKSLTG 1016
QY 327 ---KPADGSDVPNPGTT-----VGGSKQOGSSIGSI----- 354
Db 1017 QEWKP--GTVARPTTWNVLDLTGG--DSGSDVGSVLKIQAGDKIRLKSEKAAKAVI 1072
QY 355 --RVSMLL-----DDAENETASILMGFRQMIHFNFTENPDSQAQCELA 397
Db 1073 DPEVKTLILALKEGKYKTLISGKDWTPDAKSEPAVV-----KKEASPVMSAPKDELTOEIN 1127

QY 398 AQARAAKAAAGDDSAASAAALADAAQKALEAAL 426
Db 1128 AQGEKVRAAKGNKAKEVIDAEVAKILLAL 1156

Search completed: March 24, 2004, 05:57:01
Job time : 10.3108 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:47:55 ; Search time 34.0475 Seconds
(without alignments)
4188.690 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204

Sequence: 1 MWNPIGPFDITERTPAD.....QGILNALGQIASAAVVSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2204	100.0	651	16	Q927H7
2	2201	99.9	715	2	Q46166
3	723	32.8	636	16	Q824X3
4	461.5	20.9	647	16	Q84627
5	197	8.9	973	16	Q8XD04
6	193.5	8.8	2310	16	Q8CMU7
7	192.5	8.7	2055	2	Q85472
8	190	8.6	1327	2	Q9X7M2
9	190	8.6	2478	2	Q9L69
10	188	8.5	2478	2	Q9LCH2
11	186.5	8.5	790	16	Q8ZQ25
12	185.5	8.4	1822	2	Q07290
13	182	8.3	625	5	Q86RM2
14	181	8.2	2481	16	Q99QR6
15	180	8.2	621	5	Q86RM2
16	179	8.1	842	5	Q86RN1

17 176 9.0 625 5 Q86RN2
18 175 7.9 627 17 Q9HS86
19 175 7.9 855 5 Q86DH2
20 175 7.9 2345 16 Q826P6
21 173.5 7.9 2275 16 Q8NUL3
22 173.5 7.9 3360 16 Q88X86
23 173 7.8 2271 16 Q99QX4
24 172.5 7.8 812 16 Q8ZQ81
25 172.5 7.8 1128 5 Q26947
26 172 7.8 697 5 Q9NDJ0
27 172 7.8 856 5 Q86RM6
28 172 7.8 2016 5 Q9BIT0
29 170 7.7 993 16 Q8NUL3
30 169 7.7 1795 16 Q9LCJ9
31 168.5 7.6 790 16 Q8ZNI3
32 168.5 7.6 1037 10 Q7XS51
33 168.5 7.6 2232 5 Q8IFX6
34 168 7.6 993 16 Q99Q25
35 168 7.6 1545 16 Q99Q25
36 167 7.6 1033 10 Q7XVS0
37 166.5 7.6 803 5 Q86RN4
38 165.5 7.5 849 5 Q86R71
39 165.5 7.5 956 16 Q9LON7
40 165 7.5 732 5 Q86RM7
41 165 7.5 2283 2 Q8VQ99
42 164.5 7.5 639 5 Q86RM4
43 164.5 7.5 2167 2 Q845L8
44 164.5 7.5 9439 16 Q8CP76
45 164 7.4 6713 16 Q99U54

ALIGNMENTS

RESULT 1

Q927H7 PRELIMINARY; PRT; 651 AA.
ID Q927H7
AC Q927H7
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CHLPN 76 kDa HOMOLOG 1 (CT622) (Hypothetical protein).
GN CPN0728 OR CP00728 OR CP0018 OR CP0756.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=9206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138

```

RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-183;
RA Geng M.M., Schumacher A., Muehlendorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RL other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001654; AAD18867.1; -
DR EMBL; AE002165; AAF37914.1; -
DR EMBL; AP002547; BAA98935.1; -
DR EMBL; AB017159; AAP98685.1; -
DR PIR; D72042; D72042.
DR PIR; E86581; E86581.
DR PHCI-2DPAGE; Q927H7; -.
DR TIGR; CP0018; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 651 AA; 68217 MW; 47AE6C3FF2FF0F123 CRC64;

Query Match 100.0%; Score 2204; DB 16; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.3e-98;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVNPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 60
Db 1 MVNPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 60
Qy 61 LRSVAVNALMSLADKLGIGIASSNSSSTSRSDVDSTTATPTPTPTDDYKTAQTAYDT 120
Db 61 LRSVAVNALMSLADKLGIGIASSNSSSTSRSDVDSTTATPTPTPTDDYKTAQTAYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Qy 181 ASDNQAILDSLGKLTFSFDLLQAAALLOSVAANNKAAELLEMQDNVPVPGKTPAIAQSLVD 240
Db 181 ASDNQAILDSLGKLTFSFDLLQAAALLOSVAANNKAAELLEMQDNVPVPGKTPAIAQSLVD 240
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNAAGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 241 QTDATATQIEKDGNAIRDAYFAGQNAAGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Qy 301 KKFPPDSPILQAEQMWIQAQEKDLNKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 360
Db 301 KKFPPDSPILQAEQMWIQAQEKDLNKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 360
Qy 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
Db 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
Qy 421 ALEAALGKAGQQQGGIILNALGQIASAAVVSAGV 452
Db 421 ALEAALGKAGQQQGGIILNALGQIASAAVVSAGV 452

RESULT 2
Q46166 PRELIMINARY; PRT; 715 AA.
ID Q46166
AC Q46166;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 76 kDa protein.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94156481; PubMed=7509320;
RA Perez-Melgosa M., Kuo C.-C., Campbell L.;

RT "Isolation and characterization of a gene encoding a Chlamydia
RL pneumoniae 76-kilodalton protein containing a species-specific
RN epitope.";
RP Infect. Immun. 62:880-886 (1994).
RC EMBL; L23921; AAA23117.1; -.
RA PIR; I40729; I40729.
SQ SEQUENCE 715 AA; 76626 MW; 3F01C9A59DE964C5 CRC64;

Query Match 99.9%; Score 2201; DB 2; Length 715;
Best Local Similarity 99.8%; Pred. No. 9.8e-98;
Matches 451; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVNPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 60
Db 257 LVNPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDTSVERWSI 316
Qy 61 LRSVAVNALMSLADKLGIGIASSNSSSTSRSDVDSTTATPTPTPTDDYKTAQTAYDT 120
Db 317 LRSVAVNALMSLADKLGIGIASSNSSSTSRSDVDSTTATPTPTPTDDYKTAQTAYDT 376
Qy 121 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Db 377 IFTSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 436
Qy 181 ASDNQAILDSLGKLTFSFDLLQAAALLOSVAANNKAAELLEMQDNVPVPGKTPAIAQSLVD 240
Db 437 ASDNQAILDSLGKLTFSFDLLQAAALLOSVAANNKAAELLEMQDNVPVPGKTPAIAQSLVD 496
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNAAGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 497 QTDATATQIEKDGNAIRDAYFAGQNAAGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 556
Qy 301 KKFPPDSPILQAEQMWIQAQEKDLNKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 360
Db 557 KKFPPDSPILQAEQMWIQAQEKDLNKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLL 616
Qy 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 420
Db 617 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAARAAAGDSDSAAALADAQK 676
Qy 421 ALEAALGKAGQQQGGIILNALGQIASAAVVSAGV 452
Db 677 ALEAALGKAGQQQGGIILNALGQIASAAVVSAGV 708

RESULT 3
Q824X3 PRELIMINARY; PRT; 636 AA.
ID Q824X3
AC Q824X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN CCA00015.
OS Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J., Holtzaple E., Khouri H., Federova N.B., Carty H.A.,
RA Unayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavoil P.M.,
RA Fraser C.M.;
RT "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):
RL examining the role of niche-specific genes in the evolution of the
RL Chlamydiaceae.";
RL Nucleic Acids Res. 31:2134-2147 (2003).
DR EMBL; AE016994; AAP04768.1; -.
DR TIGR; CCA00015; -.
KW Hypothetical protein; Complete proteome.

```

SQ SEQUENCE 636 AA; 67020 MW; C734944DCB8CA302 CRC64;

Query Match
Best Local Similarity 38.3%; Score 723; DB 16; Length 636;
Matches 176; Conservative 84; Mismatches 166; Indels 34; Gaps 8;

QY 1 WYNTGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESK--TDSVERW 58
DB 1 WYNPV--GPIDESKINAPADLTGLMGQAANRSSEAQSITGIAGKSGSQSVSTVGR 58
QY 59 SILRSVAVNALMSLADKLGIASSNSSSSTSRSDVDSTATATPTPPPTDDYKTKQAQTA 118
DB 59 SFLSSARKSLASLEFKI---SSFFSGKIT-----PQTFDEAKTQAESAK 99
QY 119 DTIFSTSLADIQAALVSLQDVTNKKDTAETETAIAAEWETKNADAVKVGAIITELA 178
DB 100 TALQSAATTYDQPKTALQQLQDQAVKQMEQLATTDAEKATVATWKTALAEQKSLDTLNLQ 159
QY 179 KYASDNOAILDSGLKLTSPDLLQALLOSVAANNKAAELLKEMQDNPVVPKTPAIAQSL 238
DB 160 ALLTENQKLEIAKTTSSMDQIMGAAGQVETNKTABELIKOLKE-----AGVSYPVIDL 215
QY 239 VDQTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAE 298
DB 216 EKQITTSQTQVTELADAISEAAYAGKNSTAAVQQAQANNSPANIEASKTIANAKVIED 275
QY 299 AQKFPDPSPILOEAEQMWIOAEKDLKNIKPADGSDVP--NPQT--TVGSKQOQSSIGSI 354
DB 276 ALKQAPDPSPIKAAUKEQQAADILNVKPSGSDVPVIGGFGAPGSVGTQNRGATLGEV 335
QY 355 RVSMILLDDAENETASILMSGFROMIHMENTENPDSQAQQLAAQARA--AGDPSA 411
DB 336 RVSMILLTDVDNETAAILMQGFNMIDNPHQNSDFTAPLEIMNQVTLSTQINPADAEA 395
QY 412 AAALADAQKALBAALGKAGQOQOQILNALGQIASAAVVSAG 451
DB 396 TAQLQEIQOTTQDALQGTAGQDGMINALGAIITWTAASISTG 435

RESULT 4

OB4627 ID OB4627 PRELIMINARY; PRT; 647 AA.
AC OB4627;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE CHLPN 76 kDa homolog.
GN CT622.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis";
RL Science 282:754-759(1998).
DR EMBL; AE001333; AAC68226.1; -;
DR PIR; G71490; G71490.
DR PHCI-2DPAGE; OB4627; -;
KW Complete proteome.
SQ SEQUENCE 647 AA; 68525 MW; COD14C2D74473625 CRC64;

Query Match
Best Local Similarity 20.9%; Score 461.5; DB 16; Length 647;
Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;

QY 2 VNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDSDVERNSIL 61
DB 2 VNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAEAKPKESKTDSDVERNSIL 61

Db 15 MNPIINGOI-----ASNSETKEKSEEA-----SPSASSSVSSWSFL 52
QY 62 RSANVALMSLADKLGIASSNSSTSRSDVDSTTATPTPPPTDDYKTKQAQTA 120
Db 53 SSARKHALISLD--AILNKNSSPTDSLQLEASTSTSVTRVAARDYNEAKSNPDTAKSG 110
QY 121 IFTSTSLADIQAALVSLQDVTNKKDTAETETAIAAEWETKNADAVKVGAIITELAK 179
Db 111 LENATTAEYETKMAADLMAALQDMERLAKQKAEYTRIKEALQEKQ----EVIDKLNQLVK 166
QY 180 YASDNOAILDSGLKLTSPDLLQALLOSVAANNKAAELLKEMQDNPVVPKTPAIAQSL 239
Db 167 LEKQNTLKEVLTWTTSDADQIPAINSOLEINKNSADQIIKOLEGNI---SYEAVLTNAG 223
QY 240 DOTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAE 299
Db 224 EVIKASSEAGIKLGALQSIWVDAQSQAAVLAQOQNSPDNIATKGLIDAAETKVNEL 283
QY 300 QKQ---FPDPSPILOEAEQMWIOAEKDLKNIKPADGSDVPVPGTIVGSKQOQSSIGSI-- 354
Db 284 KOEHTGLTDSLVLKKAEEQISQAQKDIQEIKPS--GSDIPIVGPS--GSAASASGAVGALKS 341
QY 355 -----RVSMILLDDAENETASILMSGFROMIHMENTENPDSQAQQLAAQARA--AA 406
Db 342 SNNSGRISLILLDDVDNEMAAIAMQGFMSMIEQFVNNVPATAKELQAEQAUTAMSDQLVG 401
QY 407 GDSAAALADAQKALBAALGKAGQOQOQILNALGQIASAAVVSAG 451
Db 402 ADGELPAEIQAIKDALAQAL--KQSTDTGLATAMGQVAFVAAAKVGG 445

RESULT 5

QBXDD4 ID QBXDD4 PRELIMINARY; PRT; 973 AA.
AC QBXDD4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative membrane protein of prophage CP-933X (Putative tail fiber protein).
DE protein.
GN Z1918 OR ECS1650.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RT Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RT DNA Res. 8:11-22(2001).
DR EMBL; AE005333; AAG56007.1; -;
DR EMBL; AP002555; BAB35073.1; ALT_INIT.
DR PIR; B90835; B90835.
DR PIR; C85693; C85693.
DR GO; GO:0005198; F:structural molecule activity; IEA.

[illegible]

RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterson R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856 (2001).
 DR EMBL; AB008737; AAL19862.1; -.
 DR InterPro; IPR008969; Carboxypeptid reg.
 DR InterPro; IPR005068; Phage fiber-2.
 DR Pfam; PF03406; Phage_fiber_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 790 AA; 79050 MW; C8BE62739B80C9C3 CRC64;

Query Match 8.5%; Score 186.5; DB 16; Length 790;
 Best Local Similarity 22.9%; Pred. No. 0.33;
 Matches 113; Conservative 68; Mismatches 210; Indels 103; Gaps 18;

QY 8 GPIDETPTPPADLSAAGLEASAAKSAEQAQIAGAEAKPKESKTDVSVERWSILRSAYNA 67
 Db 91 GAMTEDDVREPEALRFEAMVEVARQASEASRNATAAGQASQATSAGQ-----AAES 144
 QY 68 LMSLDKLGIGASNSSSSTSRSDVDTATPTPPPTFDYKTAQAYDTIFTSTSL 127
 Db 145 ATAAASAAGADASATQAASSAASAESGATTK-----AGEASASAASADTART 196
 QY 128 ADIQAALVSLQAVTNIKDTATDEETALAEWETKNADAVKVGQITELAKYASDNOAI 187
 Db 197 AAASEAAKTSEANADSRSTAAGDSAAAASASATAAQTSAERAGA--SETAKTSETQA- 253
 QY 188 LDSLGLKSLFDLLQALLQSVANNKAAELKEMQDNVVPKTPA-----IAQSLVDOTD 243
 Db 254 ASSAGUAGASATAAAASEKAAAS--AAE-AKTSEITNAATSAATAAASSASASAS 310
 QY 244 ATATQIEKQGN-AIRDVAFQGNAGAVENAKSNNSINI----- 282
 Db 311 THAAASDTSASIAAQSRAAAGESATRAEEAKRAEDIADEVISLEDASLTKKGIVQLSSAT 370
 QY 283 DSAAKAIAT---AKTQAEAKKPP-DSPII-----QEA 312
 Db 371 DSDSEALASTPKAVKTVMGVOTKAPLDSPALTGTPPTPTTAAGIEIATAFVAARV 430
 QY 313 EQMVIQAEK---DLKNIKPADGSDVNPPTGVGSKQGGSGIGSRVSMLLDDAEINETS 369
 Db 431 AOLVGAPEALDTLKEALADLGN-D-PNFATIV-----LNKLAGKPLDITLALSG 480
 QY 370 ILMGFRQMIHMFNTENPDSSQAQOELAAQAPAAKAGDSDAAALADAQKALEALGKA 429
 Db 481 KSVDPGLIEYIGLRNTID-----KAAGALPAGGTAVANRLVSRGALPALTGTT 528
 QY 430 -GQQQGIILNALGQI 442
 Db 529 RGSDSGLI--MGEV 540

RESULT 12
 Q07290 PRELIMINARY; PRT; 1822 AA.
 AC Q07290;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Ept* protein.
 GN Ept*.
 OS Streptococcus suis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1307;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=1890;
 RX MEDLINE=93328288; PubMed=8335363;
 RA Smith H.E., Reek F.H., Vecht U., Gielkens A.L.J., Smits M.A.;
 RT "Repeats in an extracellular protein of weakly pathogenic strains of

RT Streptococcus suis type 2 are absent in pathogenic strains.";
 RL Infect. Immun. 61:3318-3326 (1993).
 DR EMBL; X71880; CAA50714.1; -.
 DR PIR; S33441; S33441.
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR005877; Gpos_YSIK.
 DR Pfam; PF04650; YSIK_signal; 1.
 DR TIGRFAMS; TIGR01168; YSIK_signal; 1.
 SQ SEQUENCE 1822 AA; 192632 MW; 3838960C77641D7D CRC64;

Query Match 8.4%; Score 185.5; DB 2; Length 1822;
 Best Local Similarity 22.8%; Pred. No. 1;
 Matches 119; Conservative 88; Mismatches 207; Indels 107; Gaps 22;

QY 10 IDETERTPPADLSAAGLEASAAKSAEQAQIAGAEAKPKESKTDVS---ERWSILRSAYN 66
 Db 1072 IDDNPNLTPE-----KESAKNAVEEAQVATA-AIDKASTPDPAQVQVEEDKGV--A 1122
 QY 67 ALMSLDKLGIGASNSSSSTSRSDVDTATPTPPPTFDYKTAQAYDTIFTSTSL 112
 Db 1123 LITAKADAKGVIAAKLADEIKKLEDKQAEAKAIDASTMTNEEKAIAKALQDVVVGKA 1182
 QY 113 Q---AQTAYDTIFTSTSLADIQAALV---SLQDAVTNIKDTATDEETALAE-WETK 163
 Db 1183 ELEDAARVATNEIHEATITTEKAKAEAGEKSLDTGTGKEARDAVELAKDKELAKEIRTE 1242
 QY 164 NADAVKVGQAITTELAKYA-----SDNOA-----ILDSLGKLTSPDL 199
 Db 1243 EEEATKIVEKLAEDTRKAIEDNPNLSDEDKQAEIKCLTDVAVAKTLATIRDNADKRTQ--- 1299
 QY 200 LQALLQSVANNKAAELKEMQDNVVPKTPAIAQSLVDOTATA-TQIEKGNALRD 258
 Db 1300 -EAKQALADLEKAKTQK-IADKAAIDELTILVKDGELEATKQDAKNTAKDAKAAAKE 1357
 QY 259 AYPAGQVAGSAG-----VENAKSNNSIS-----NIDSAKAAIA----- 290
 Db 1358 AIASNPNTLDAEKTFTDVADEAVAKANDASATSADVQKEEDAGVAALAEVDLRAK 1417
 QY 291 -TAKTQIAE---AQKFPDPSPILOEAFQWVIAEKLNKIKPADGSDVNPPTGVGSK 345
 Db 1418 QDAKNKIAKDAKAAAEKAGSNPNLTDAEKTFTDVADEAVAKANDASATSADV--QK 1475
 QY 346 QQSGSIGSIRVSM---LDDAENETASILMSGFPMHMTENTENPDSSQAQOELAAQARA 402
 Db 1476 EEDAGVAALAEVDLRAKQDAKNTAKESDAKSAIDANPNLTDAEKESAKKAVDADAKA 1535
 QY 403 AKAGDSDSAAALADAQKALEAALGKAGQQGQGIILNALGQIA 443
 Db 1536 ATDAID--ASTSPVEAQSAEDKGVGSIADQ--VLDAAKQDA 1572

RESULT 13
 Q06RN6 PRELIMINARY; PRT; 625 AA.
 AC Q06RN6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Merozoite surface protein 3 alpha (Fragment).
 OS Plasmodium vivax.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5855;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=IRC-Brazil 67;
 RX MEDLINE=22346636; PubMed=12458823;
 RA Rayner J.C., Corredor V., Feldman D., Ingravallo P., Iderabdullah F.,
 RA Galinski M.R., Barnwell J.W.;
 RT "Extensive polymorphism in the plasmodium vivax merozoite surface coat
 protein MSP-3alpha is limited to specific domains.";
 RL Parasitology 125:393-405 (2002).
 DR EMBL; AF491946; AAC020877.1; -.
 FT NON_TER 1 1

```
FT NON TER 625 625
SQ SEQUENCE 625 AA; 65303 MW; ADA71E17C2914047 CRC64;

Query Match 8.3%; Score 182; DB 5; Length 625;
Best Local Similarity 21.6%; Pred. No. 0.41; Indels 82; Gaps 12;
Matches 102; Conservative 66; Mismatches 223;

Qy 23 AQLGSAANKSABAQRIA-GAEAKPKESKTDVSVE-----RWSILRS 63
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 81 AEEAABAKNQAAETAKAADAEEAAGKENDLDDVKSQVKIAEEASTNAKKNKTEAIVE 140
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 64 AVNALMSL-----ADKLGIASNSSSSSTSSADVDSTATAPPPPPDDFYKTOA 114
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 141 IVKAVVAKEEAQKASDEAQACEKAQKAHAKQASDTTKTVERFKTNAEAAAKNAKEKA 200
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 115 QTAVDYTFSTSLADIQALVSLQDAVTNIDKDTAATDETAIAAEWETKNADAVKGAQI 174
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 201 ENANKATEASANELSVAKQKATDAETAADKAKKEQVKAELAE-----VAKAKVAKEE 255
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 175 TELA-KYASDNOAILDSLGKLTSPDLLQAALLQSVANNKAAEILLKEMQDNPPVPGKTPA 233
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 256 ADAQKKAEEAKKIVDKIADKQSEVPEAQKAAEFATETVKATTAATEAGKNAQEAESPE 315
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 234 IAQSLVDOTDATAQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAIATAK 293
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 316 EAEEKAETSDAVKQKAD-----AAEKAAGEAKKASIETEI-----AVEVAK 356
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 294 TOIAEAKKFFPDSPILOAEQWIOAEKDLNKPFA-----DGSVDPNPGTIVGSK 345
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 357 AEVLNAEVK-----KTAQAEKQATEAKQAEKAKAAAEAEKTHEKAEKVGESTKASDX 412
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 346 QQGSISGIRVSMLLDDAENETASILMSGFRQIMHFNTEPNDSQAQAEILAAQARAKA 405
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 413 AQQENKNAKAS---EEAENRAVDALAEYAVEAHLARTKNAEASAKSATDMSLEKAKE 469
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 406 AGDDSA-----AALADAQKALEALGKAGQOQOILNALGQIAAA 446
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 470 BAIDAAANIAHQWLKATQATIAKEK--EAAKVAEKAQTAANVVKDKAKA 520
Db : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q99QR6 PRELIMINARY; PRT; 2481 AA.
AC Q99QR6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE FmtB protein (OR FMTB OR SAV2160 OR SA1964.
GN FMTB (MRP) OR FMTB OR SAV2160 OR SA1964.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPICIRS-S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa K., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58322.1; -.
DR EMBL; AP003136; BAB43253.1; -.
DR PIR; D90011; D90011.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
```

```

RA Rayner J.C., Corredor V., Feldman D., Ingravallo P., Iderabullah F.,
RA Galinski M.R., Barnwell J.W.;
RT "extensive polymorphism in the plasmodium vivax merozoite surface coat
RT protein MSP-3alpha is limited to specific domains.";
RL Parasitology 125:393-405(2002).
DR EMBL: AF491961; AAC020891.1; -.
FT NON_TER 1 621
FT NON_TER 621 621
SQ SEQUENCE 621 AA; 64825 MW; BFD2CC64CDB29AA6 CRC64;

Query Match 8.2%; Score 180; DB 5; Length 621;
Best Local Similarity 22.3%; Pred. No. 0.51;
Matches 105; Conservative 68; Mismatches 217; Indels 80; Gaps 14;

QY 23 AQGLEASAANKSABAQRIA-GAEAKPKESKTDVSVERNSILRSVAVNLMSLADKLGIA--- 78
Db 81 AEEAAAEAKQAAIAEKAADAADAAEAKENKLLDVK--SQVQIAVEASTKAKDKKTEAEIA 138
QY 79 -----SSNSSSTSRSDVSDTSTAATPPPTDDVKYTOACTAYDTFT 123
Db 139 VEIVKAVVAEEAQKADEQACEKAKAHAKAKASDITKTIVETFKTNHAAANAKVD 198
QY 124 STSLADIQAALVSLQDVAITNIKOTATDEETAIAAEWETKNAD-----AVKVGQAQITEL 177
Db 199 EAGKAETAATDAKSTEDLSVAKDKAKAAETAIEAKKQVQKAEATAAEVAKAKVAKEBADA 258
QY 178 A-KYASDQALLDSLGKLTSPDLLQALLOSVAANNKKAELLKEMQNDPVVPGKTPAIAG 236
Db 259 AQKAAEAKKIVDKIAKSDNVPEAKAA-----EFATETVTKATTAATEAGKN---AQ 308
QY 237 SLVDQTDATATQIEKGNNAIRDVAFAGNAGSAGVENAKSNNSIKNIDSKAAAIATAKTQI 296
Db 309 EAAKSPBEAAT-----SDAVKGKADAAEKAAGEAKKAMIEI-----AIEVAKAEV 355
QY 297 AEAKKPFDPFILOAEQMWIOAKDLINKIPA-----DGSDEVNPGTTVGGSKQOQ 348
Db 356 IYAEVK---KTAQEAESKDATEKEQEKAKAAAAEAAKTHGEKAERGVGESTKKAHSDEAQ 411
QY 349 SSIGSIRVSMLLDDAENETASILMSGFRQIMHMENTENPDPSQAAQQLAQAARAKAAGD 408
Db 412 ENKVAKADAS---EFAENRAVDAAEEAVAEVHAHARTKNVAESAASATMDMELEKAKEAI 468
QY 409 DSA-----AAALDAOKALBAALGKAGQOQQGILNALGQIIASAA 446
Db 469 DAANIAHQKWKATQAATIAKEK--EAAKVAAEKATTAANVVVKDKKAKA 516

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:53:56 ; Search time 12.3321 Seconds
(without alignments)
3525.629 Million cell updates/sec

Title: US-10-608-559-6

Perfect score: 2204
Sequence: 1 MWNPIGPGIDETERTPPAD.....QGILNALQIAAAVVSAGV 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.78.*

- 1: Pirl.1.*
- 2: Pirl.2.*
- 3: Pirl.3.*
- 4: Pirl.4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2204	100.0	651	2	D72042 conserved hypothet
2	2204	100.0	651	2	CHLPN 76 kDa homol
3	2201	99.9	715	2	hypothetical 76K p
4	461.5	20.9	647	2	hypothetical prote
5	197	8.9	971	2	probable tail fibe
6	197	8.9	973	2	probable membrane
7	192.5	8.7	2055	2	extracellular matr
8	185.5	8.4	1822	2	EF protein - strep
9	184	8.3	1122	2	probable tail fibe
10	181	8.2	2481	2	FmtB protein [impo
11	176.5	8.0	545	2	Htr7 transducer [i
12	176.5	8.0	545	2	halobacterial tran
13	175	7.9	627	2	Htr14 transducer [
14	173	7.8	2271	2	hypothetical prote
15	172.5	7.8	1128	2	R27-2 protein - Tr
16	168.5	7.6	2232	2	hypothetical prote
17	168	7.6	993	2	hypothetical prote
18	164	7.4	5713	2	hypothetical prote
19	163.5	7.4	2155	2	conserved hypothet
20	163.5	7.4	2155	2	hypothetical prote
21	160.5	7.3	544	2	transducer protein
22	160	7.3	536	1	A47190 transducer [i
23	160	7.3	536	2	Htr1 transducer [i
24	158.5	7.2	4776	2	cell wall surface
25	157	7.1	641	2	methy-accepting c
26	156	7.1	810	2	halobacterial tran
27	156	7.1	810	2	Htr5 transducer [i
28	156	7.1	2186	2	hypothetical prote
29	154.5	7.0	1238	2	probable exonuclea

ALIGNMENTS

RESULT 1

D72042
Conserved hypothetical protein CP0018 [imported] - Chlamydia pneumoniae (strains CWI
N; Alternate names: chlpn 76 kDa homolog 1 (ct622); hypothetical protein CPN0728
C; Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C; Date: 23-Apr-1999 #sequence_revision-23-Apr-1999 #text_change 11-May-2000
C; Accession: D72042; D81523
R; Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A; Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A; Reference number: A72000; MUID:99206606; PMID:1092388
A; Accession: D72042
A; Molecule type: DNA
A; Residues: 1-651 <ARN>
A; Cross-references: GB:AE001654; GB:AE001363; NID:94377031; PIDN:AAD18867.1; PID:9437703
A; Experimental source: strain CWL029
R; Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A; Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A; Reference number: A81500; MUID:20150255; PMID:10684935
A; Accession: D81623
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-651 <REA>
A; Cross-references: GB:AE002165; GB:AE002161; NID:97188948; PIDN:AAF37914.1; PID:9718894
A; Experimental source: strain AR39, HL cells
C; Comment: This sequence was originally identified as homologous to part of a sequence
PIR:H71490.
C; Genetics:
A; Gene: CPN0728; CP0018

30 154 7.0 5327 2 T13564 microtubule-associ
31 153.5 7.0 1147 2 T35781 hypothetical prote
32 153 6.9 860 2 T14650 tail fiber protein
33 153 6.9 860 2 T14650 phage lambda-relat
34 153 6.9 1365 2 T30822 lmp1 protein - Myc
35 153 6.9 3624 2 AD0835 large repetitive p
36 152 6.9 892 2 T50385 related to transcr
37 150 6.8 990 2 T51618 nucleolar phosphop
38 150 6.8 1430 2 AF0351 probable autotrans
39 149.5 6.8 582 2 S24545 intermediate filam
40 149 6.8 778 2 T48897 transducer protein
41 149 6.8 1156 2 T34852 probable secreted
42 149 6.8 1561 1 S06839 surface antigen sp
43 148.5 6.7 764 2 A84328 Htr2 transducer [i
44 148.5 6.7 765 1 T44946 transducer protein
45 148.5 6.7 3890 2 C89921 hypothetical prote

Query Match 100.0%; Score 2204; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 7.2e-110;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAAGAEKPKESKTSVERWSI 60
Db 1 MWNPIGPGIDETERTPPADLSAQGLEASAAKSAQAQIAAGAEKPKESKTSVERWSI 60
Qy 61 LRSVNALMSLADKLGITASSNSSSTSRSDVSTTATATPPPTDDYKQATAYDT 120
Db 61 LRSVNALMSLADKLGITASSNSSSTSRSDVSTTATATPPPTDDYKQATAYDT 120
Qy 121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 180
Db 121 IFSTSLADIQAALVSLQDAVTNFKDTAATDEETAIAAEWETKNADAVKGAQITELAKY 180
Qy 181 ASDNQAILDSGLKITSFDLLQAALQSVANNKAAELLKEMQDNFVVPVPGKTPAQSLVD 240
Db 181 ASDNQAILDSGLKITSFDLLQAALQSVANNKAAELLKEMQDNFVVPVPGKTPAQSLVD 240

```

QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
QY 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVNPNGTTVGGSKQGGSSIGSIRVSMML 360
DB 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVNPNGTTVGGSKQGGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAAGDSDSAAALADAQK 420
DB 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAAGDSDSAAALADAQK 420
QY 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452
DB 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452

RESULT 2
E86581
ChLPN 76 kDa homolog 1 (CT622) [imported] - Chlamydothila pneumoniae (strain J138)
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: E86581
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Takuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise, N.
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: E86581
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-651 <RES>
A:Cross-references: GB:BA000008; NID:g8979100; PIDN:BAA98935.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: CPj0728

```

```

Query Match 100.0%; Score 2204; DB 2; Length 651;
Best Local Similarity 100.0%; Pred. No. 7, 2e-110;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 60
DB 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 60
QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATPTPPPTFDYKTAQATAYDT 120
DB 61 LRSVAVNALMSLADKLGIASSNSSSTSRGADVSTTATPTPPPTFDYKTAQATAYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAETETAIAEWETKNADAVKVGAIITELAKY 180
DB 121 IFTSTSLADIQAALVSLQDAVTNIKDTAETETAIAEWETKNADAVKVGAIITELAKY 180
QY 181 ASDNQAILDSGLKTSFDLLQALLOSVAANNKAELLKEMQDNVVPKTPAIAQSLVD 240
DB 181 ASDNQAILDSGLKTSFDLLQALLOSVAANNKAELLKEMQDNVVPKTPAIAQSLVD 240
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
DB 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
QY 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVNPNGTTVGGSKQGGSSIGSIRVSMML 360
DB 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVNPNGTTVGGSKQGGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAAGDSDSAAALADAQK 420
DB 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAAGDSDSAAALADAQK 420
QY 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452
DB 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452

```

```

RESULT 3
140729
hypothetical 76k protein - Chlamydothila pneumoniae (strain AR39)
C:Species: Chlamydothila pneumoniae, Chlamydia pneumoniae
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
C:Accession: 140729
R:Perez-Melgosa, M.; Kuo, C.
Infect. Immun. 62, 880-886, 1994
A:Title: Isolation and characterization of a gene encoding a Chlamydia pneumoniae 76-kDa
A:Reference number: 140729; MUID:94156481; PMID:7509320
A:Accession: 140729
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-715 <RES>
A:Cross-references: GB:L23921; NID:g435961; PIDN:AAA23117.1; PID:g435962
A:Experimental source: strain AR-39
C:Comment: This is the hypothetical translation of a sequence that was reported as two

```

```

Query Match 99.9%; Score 2201; DB 2; Length 715;
Best Local Similarity 99.8%; Pred. No. 1, 2e-109;
Matches 451; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 60
DB 257 LWNPIGPGPIDETERTPPADLSAQGLEASAAKSAEAOIAGAEAKPKESKTDVSVERWSI 316
QY 61 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATAYDT 120
DB 317 LRSVAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTFDYKTAQATAYDT 376
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAETETAIAEWETKNADAVKVGAIITELAKY 180
DB 377 IFTSTSLADIQAALVSLQDAVTNIKDTAETETAIAEWETKNADAVKVGAIITELAKY 436
QY 181 ASDNQAILDSGLKTSFDLLQALLOSVAANNKAELLKEMQDNVVPKTPAIAQSLVD 240
DB 437 ASDNQAILDSGLKTSFDLLQALLOSVAANNKAELLKEMQDNVVPKTPAIAQSLVD 496
QY 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
DB 497 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 556
QY 301 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVNPNGTTVGGSKQGGSSIGSIRVSMML 360
DB 557 KKFPPSPILQEAQWVQIAEKDLKNIKPADGSDVNPNGTTVGGSKQGGSSIGSIRVSMML 616
QY 361 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAAGDSDSAAALADAQK 420
DB 617 DDAENETASILMSGFQMIHNTENPDSSQAQOELAAQARAAGDSDSAAALADAQK 676
QY 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452
DB 677 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 708

```

```

RESULT 4
G71490
hypothetical protein CT622 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
N:Alternate names: chlpn 76kda homolog CT622
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: G71490
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:9900809; PMID:9784136
A:Accession: G71490
A:Molecule type: DNA
A:Residues: 1-647 <ARN>
A:Cross-references: GB:AE001333; GB:AE001273; NID:g3329068; PIDN:AAC68226.1; PID:g33290
A:Experimental source: serotype D, strain UW-3/Cx
C:Comment: This sequence was originally identified as homologous to part of a sequence
PIR:E72042).

```

C;Genetics:
A;Gene: CT622

Query Match 20.9%; Score 461.5; DB 2; Length 647;
Best Local Similarity 29.7%; Pred. No. 2.6e-17;
Matches 138; Conservative 86; Mismatches 192; Indels 49; Gaps 13;

QY 2 VNPFGPDPIDETERTPPADLSAAGLEASAAKAGAEQRIAGABAKPKESKTDVSVERWSIL 61
DB 15 MNPLINGQI-----ANSETKESKSEA-----SPSASSSVSSWSPL 52

QY 62 RSVAVALMSLADKLGIIASSNSSSTSRs-ADVSTTATATPPPPFDYKTAQATYDT 120
DB 53 SSAKHALISLRD--AILNKNSSPTDLSQLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110

QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEE-TAJAABWETKNADAVKVGQAQITELAK 179
DB 111 LENATTIAEYETKADLMAALQDMERLAKOKAEVTRIKEALQEKQ-----EVIDKLNQLVK 166

QY 180 YASDNQAILDSLGKLTSTFDLLQALLOSVAANNKKAELLKEMDNPNVVPKTPALAQSLV 239
DB 167 LEQNQNTLKETITTTDSADQIPALINSQLEINKNSADQIIKDLEQONI---SYEAVLTNAG 223

QY 240 DQTDATATQTEKQGNARTRAYFACQNASGAVENAKSNNSISNIDSAKAAATATKTQIAEA 299
DB 224 EVIKASSEAGIKLQALQSLVDAGDSQAAVLQACQNNSPDNIAATKLLIDAAETKVNEL 283

QY 300 QKK---FPDPSILOAEQOMVIOAEKDLNLIKPADGSDVPNPGTTVGGSKQGGSSIGSI-- 354
DB 284 KOEHTGLTDSPLVKABEQISOAKDIOETKPS-GSDPIVGPSS-GSAAGSAGVAGLKS 341

QY 355 -----RVSMILLDDAENETASTILMSGFRQMIHMFENTENPDQAAQOELAAQARAAK---AA 406
DB 342 SNNSGRISLILLDDVNEMAIAMGFRSMIBQFNVPNPATAKELQAMEAQITAMSDQLVG 401

QY 407 GDSAAALADAQXALFAALGKAGCQOQGIINALGOIASAAVVSAG 451
DB 402 ADGELPAEIQAKDALAQAL-KQFSTDGLATAMGQVAFAAKVG 445

RESULT 5

B90835
probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: B90835
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90835
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-971 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA035073.1; PID:g13361114; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1650

Query Match 8.9%; Score 197; DB 2; Length 971;
Best Local Similarity 22.9%; Pred. No. 0.0049;
Matches 119; Conservative 63; Mismatches 208; Indels 130; Gaps 14;

QY 26 LEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSILRSVAVNMLSLADKLGIIASSNSSSS 85
DB 110 VEEVARNASAVAQNTAAAKKSADASTSAREAAHTATDAADSARAASSTAGQAAS-SAQ 168

QY 86 TSSRADVDSTTATPTPPPPFDYKTAQATYDTFTSTSLADIQAALVSLQDAVTNIK 145
DB 169 ASSAGTASTKATEASKAAAESEKSAATISAGAAKTSET-----NAAVQQAASATAS 223

QY 146 DTAATDEETAIAAEWETKNADAVKVGQAQITELAKYASDNQAILDSLGKLTSTFDLLQAL 205

DB 224 TATTKASEAASA-----RDASAKAEAKSSETSAASSASSAASSATAAGNS-----AKAA 274

QY 206 QSVANNKKAELLKEMDNPNVVPKTPA-----IAQSLVDQTDATATQIEKD-----252

DB 275 KTSETNAKSETAEQAASASAAAGSKTAAALSASAASTAGQASASATAGAKSAESAASA 334

QY 253 -----GNAIRDYFAGQNASGA-----VENAKSNNSIS-----280

DB 335 STATTKAGEATEQASAAASASAAAKTSETNAKASETSABESKTTAAASSASSASSASSAS 394

QY 281 -----NIDSKAAIATAKTQIAE-----AQKKFPDPSILOAEQOMVIOAEKDLK 324

DB 395 ASKDEATROQASAAKSATTAASKATKAEEAASATAAQSKSTAESAATRAETAAKRAEDIAS 454

QY 325 NIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMILLDDAENETASTILMSG-----374

DB 455 AVALEDAS-----TTKKGIVQLSSATNSTSESLAATPKAVKAAVELANGKYTAQDATTA 508

QY 375 FROMHMFENTENPDQAAQOELAAQARAACAAGD-----D 409

DB 509 QKGIVQLSNATNSTSEM---LAATPKSKVAAVYDLANGKYTAQDATTAQKGIVOLSSATN 564

QY 410 SARAALADAKALEALGKAGCQOQGIINALGOIASAAVVS 449

DB 565 SASSETLAATPKAVKAAND-----NANGRVPSARKVN 595

RESULT 6

C85693
probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (strai
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: C85693
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85693
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-973 <STO>
A;Cross-references: GB:AE005174; NID:g12514847; PIDN:AAG56007.1; GSPDB:GN00145; UWGP:Z19
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1918

Query Match 8.9%; Score 197; DB 2; Length 973;
Best Local Similarity 22.9%; Pred. No. 0.0049;
Matches 119; Conservative 63; Mismatches 208; Indels 130; Gaps 14;

QY 26 LEASAANKSAEQAQRIAGAEAKPKESKTDVSVERWSILRSVAVNMLSLADKLGIIASSNSSSS 85
DB 112 VEEVARNASAVAQNTAAAKKSADASTSAREAAHTATDAADSARAASSTAGQAAS-SAQ 170

QY 86 TSSRADVDSTTATPTPPPPFDYKTAQATYDTFTSTSLADIQAALVSLQDAVTNIK 145
DB 171 ASSAGTASTKATEASKAAAESEKSAATISAGAAKTSET-----NAAVQQAASATAS 225

QY 146 DTAATDEETAIAAEWETKNADAVKVGQAQITELAKYASDNQAILDSLGKLTSTFDLLQAL 205

DB 226 TATTKASEAASA-----RDASAKAEAKSSETSAASSASSAASSATAAGNS-----AKAA 276

QY 206 QSVANNKKAELLKEMDNPNVVPKTPA-----IAQSLVDQTDATATQIEKD-----252

DB 277 KTSETNAKSETAEQAASASAAAGSKTAAALSASAASTAGQASASATAGAKSAESAASA 336

QY 253 -----GNAIRDYFAGQNASGA-----VENAKSNNSIS-----280

DB 337 STATTKAGEATEQASAAASASAAAKTSETNAKASETSABESKTTAAASSASSASSASSAS 396

QY 281 -----NIDSKAAIATAKTQIAE-----AQKKFPDPSILOAEQOMVIOAEKDLK 324

Db 397 ASKDEATQASAKSATTASTKATEAAGSATAAQAQSKSTAESAATRAETAAKRAEDIAS 456
 QY 325 NIKPADGSDVNPFTVGGSKQGGSGISIRVSMLLDADENETASILMSG----- 374
 Db 457 AVALEDAS-----TTKKGIVQLSSATNSTSESAAATPKAVKAAVELANGKYTAQADATTA 510
 QY 375 FQMIHMENTENPDSCAAQQLAAQARAAGD-----D 409
 Db 511 QGGIVQLSSATNSTSEM-----LAATPKSVKAAAYDIANGKYTAQADATTAQKGVQLSSATN 566
 QY 410 SAAALADAQAALKAALGAGQGGQGGILNALGQIAGAAVVS 449
 Db 567 SASSETLAATPKAVKAAND-----NANGRVPSARKVN 597

RESULT 7
 T31110
 extracellular matrix binding protein - Abiotrophia defectiva (fragment)
 C:Species: Abiotrophia defectiva
 C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C:Accession: T31110
 R:Manganelli, R.; van de Rijn, I.
 Infect. Immun. 67, 50-56, 1999
 A:Title: Cloning and characterization of emb, a gene encoding the major adhesin of Streptococcus
 A:Reference number: Z20988; MUID:99081722; PMID:9864195
 A:Accession: T31110
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2055 <MAN>
 A:Cross-references: EMBL:AF067776; NID:g3249002; PID:g3249003; PIDN:AAD03320.1
 C:Genetics:
 A:Gene: emb

Query Match 8.7%; Score 192.5; DB 2; Length 2055;
 Best Local Similarity 23.2%; Pred. No. 0.022;
 Matches 110; Conservative 77; Mismatches 187; Indels 101; Gaps 21;

QY 19 ADLSAQGLEASANKSAEQAQRIAGABA-----KPKESKTSVERMSILRSVAKNALMSLD 73
 Db 1371 ATKAKNAIDATSNDETAKQNEGTAQINAVPTPKAKTD-----AKNAVTAQAD 1420
 QY 74 KLGIASNSGSSST-----SRSDVSTTATATPTPTPTDDYKTQATYDTFTSTSLA 128
 Db 1421 RKKDAIENDPNLTREBKVAAKAKVDA-----EAKKAKDAIDAATNSA 1462
 QY 129 DIOAALVSLQDAVNIKD--TAATDEETATAAEWETKNADAVKVGAGITELAKYASDNOA 186
 Db 1463 DVTAQNEGKTAINDVPQPTTAKTDKNAVTAQADAKK--DAIEKDPNLTREBKDA--KA 1519
 QY 187 ILDSLGLKLTSLDLOAALLQS--VANNKAAEALLKEMQDNPPVPGKTPAIAQSLVDQTD 244
 Db 1520 KVDAEAKKAK--DAIDAATSNADVTAQNEGKTAINDVPQPTT--AKTDA--KNAVTAQADA 1575
 QY 245 TATQIEKDGNAIRDAFA-----GONASGAVENAKSN-----NSISNI 282
 Db 1576 KDAIEKDPNLTREBKDAKAKVDAEAKKAKDAIDAATSNADVTAQADKNAINAVPQT 1635
 QY 283 DSAAKATATKTOIABAQK--KFPDSPILQEAQWVIOAEKDLKNIKPADGSDVNPQTTV 341
 Db 1636 PTAKTDKNAVTAQADAKDAIENDANLTREBKDAKAKVDAEATKAKNAIDAATSNADV 1695
 QY 342 GGSQKQGS--SIGSTI--RVSMILLDAEN---ETASILMSGFRQMIHMFNTEN-----PDSQ 390
 Db 1696 TAKNEGKTAINDVPQPTTAKTDAKNAVDAQATDKKSAI-----ENDPALTREKD 1746
 QY 391 AAQCELAQAAPAKAAGDGSAAALADAQKALEALGKAGCQGGQILNALGQIASA 445
 Db 1747 AAKAKVDAEATKAKNAIDAATSNADVTAQK-----DAG--KNAINAVPQTTA 1792

RESULT 8
 S33441

EF protein - Streptococcus suis
 C:Species: Streptococcus suis
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
 C:Accession: S33441
 R:Smith, H.E.; Reek, F.H.; Vecht, U.; Gielkens, A.L.J.; Smits, M.A.
 submitted to the EMBL Data Library, May 1993
 A:Description: Repeats in an extracellular protein of wek-pathogenic strains are absent
 A:Reference number: S33441
 A:Accession: S33441
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1822 <SMI>
 A:Cross-references: EMBL:X71880; NID:g298031; PIDN:CAA50714.1; PID:g298032

Query Match 8.4%; Score 185.5; DB 2; Length 1822;
 Best Local Similarity 22.8%; Pred. No. 0.045;
 Matches 119; Conservative 88; Mismatches 207; Indels 107; Gaps 22;

QY 10 IDEVETPPADLSAQGLEASANKSAEQAQRIAGAEKPKESKTSV---ERWSILRSVAVN 66
 Db 1072 IDDNPNLTPE-----KESAKNAVEEAQKAVATA-AIDKASTPDVAVQVEEDKGV--AAIN 1122
 QY 67 ALMSLADKLGIASSNSSSSSTSRSD-----VDSTTAT-----APTPTPTFDYKT 112
 Db 1123 LITAKADAKGVIAKLADEIKLEDKQAEAKAIDASTMTNEEKAIKALQDVVDKGA 1182
 QY 113 Q-----AQPAYDTFTSTSLADIQAALV---SLQDAVTNIKDTAATDBETAIAAE--WETK 163
 Db 1183 ELEDAARVATNEIHEATTTEKAKAAELAGEKSLTDTGKEARDAVELAKDELAKEAIRTE 1242
 QY 164 NADAVKVGAGITELAKYA-----SDNOA-----ILDSLGLKLTSLD 199
 Db 1243 EEEATKIVEKLAEDTRKAIEDNPNLSDEDKQAEIKKLTDAVAKTLATIRDNADKKQ--- 1299
 QY 200 LQAALLQSVANNKAAEALLKEMQDNPPVPGKTPAIAQSLVDQTDATA--TQIEKDGNAIRD 258
 Db 1300 -EAEKAAALADLEKAKETQK--IADKAAIDRLTILVKDGELEATKQDAKNKIADKAAAAKE 1357
 QY 259 AVFAGQNASGA-----VENAKSNNSIS-----NIDSAKAAIA----- 290
 Db 1358 AIASPNLTDAEKKTFTDAVDAEAVAKANDASATSPADVQKEEDAGVAAIAEDVLDAAK 1417
 QY 291 -TAKTQIAE-----AQKKFPDSPILQEAQWVIOAEKDLKNIKPADGSDVNPQTTVGGSK 345
 Db 1418 QDAKNKIADKAAAAKEATGSPNLTDAEKKTFTDAVDAEAVAKANDASATSPADV--QK 1475
 QY 346 QCGSSIGSIRVSM---LDDAENETASILMSGFRQMIHMFNTENPDSCAAQQLAAQARA 402
 Db 1476 EEDAGVAAIAEDVLDAAKDAKNAKSAKSAIDANPNLTDAEKESAKKAVDADAKA 1535
 QY 403 AKAAAGDGSAAALADAQKALEALGKAGCQGGQILNALGQIA 443
 Db 1536 ATDAID--ASTSPVEAQSAEDKGVGSIAD--VLDAAKQDA 1572

RESULT 9
 G64887
 probable tail fiber protein QP37 - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
 C:Accession: G64887; T09189
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: G64887
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1122 <BLAT>
 A:Cross-references: GB:AE000234; GB:U00096; NID:g1787633; PIDN:AACT4454.1; PID:g1787636
 A:Experimental source: strain K-12, substrain MG1655
 R:Aiba, H.; Baba, T.; Fujita, K.; Hayashi, K.; Inada, T.; Isono, K.; Itoh, T.; Kasai, H.

A;Accession: D00111
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2481 <KUR>
 A;Cross-references: GB:BA000018; PID:g13701961; PIDN:BA043253.1; GSPDB:GN00149
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: fntB(mrp)
 Query Match 8.2%; Score 181; DB 2; Length 2481;
 Best Local Similarity 23.2%; Pred.No. 0.12;
 Matches 127; Conservative 91; Mismatches 204; Indels 126; Gaps 25;

```

Query Match      8.0%; Score 176.5; DB 2; Length 545;
Best Local Similarity 21.4%; Pred. No. 0.029;
Matches 98; Conservative 79; Mismatches 195; Indels 87; Gaps 17;

QY 12 ETRTPPADLSAOGLEASAAKSAEQRAGAPAKPKESKTDSEVRWSILRSVNAWMLSL 71
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 118 EEERA-----EAEERAREKAEQQAER---QTAAEASAKQDARERSAEIEQAADLESQ 169

QY 72 ADKGLGIASNSGSGSRSDVGTATTATPTPPPT-FDDYKTKQAQYNDTI-----FTST 125

```


Db 170 ATEVG-ATLEAASDGLTARVDATTDAENAEIAEVATVNDMLTMTERTIDEIQGFSTNVTI 228
QY 126 SLADIQAALVSLQDVTNKTOT-----KATDEE-----TAAAEWETKNADAVKGAQITE 176
Db 229 ASRETAGAKETQDSQTSVQIEAGTDQDREQLSVAEEMDSYSATVEEVAATAQS 289
QY 177 LAKYASDNQALIDSGLKIT-----SFDLLQAALQSVANNKAAELKEMQDNVPVPGK 230
Db 289 VADTAADTTDVA-TAGQTAEDDAIDAIDAVQETMTQTVANVDALEDLTTEIDD----- 340
QY 231 TPAIAQSLVQDTATATQIEKQGNIRDAYFAGQNASG-----AVENAKNSNSI 279
Db 341 IAEILISDIAEQTNMLA--LNANIEAARAGSGGSGDGFVAVADEVKELATESORSAXDI 398
QY 280 SN-IDSAXAAIATAKTOIAEAKKPPDSPILOAEQMVIAEQLKLNKIPADGSDVPNPG 338
Db 399 AELIEVQSQATTVEEI-----ISQAMDEQAQORSERVSS-----VDDIATISQATAD 445
QY 339 TTVGSGKQGGSIGIRVSMLLDDAENETASIMSGFROMIHMFNTENPDSPQAQQELAA 398
Db 446 ETTDGVQE-----ISQAMDEQAQORSERVSS-----VDDIATISQATAD 484
QY 399 QARAKAGDSDAAALADAQAALGALGKAGQOQGIIN 437
Db 485 RAENVSAASEEQ-TASITEVTSLSQSLAAQADLTLEDRIN 522
RESULT 12
T46811
Halobacterial transducer protein V [imported] - Halobacterium salinarum
C:Species: Halobacterium salinarum
C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 15-Sep-2000
C:Accession: T46811
R:Rudolph, J.; Nordmann, B.; Storch, K.F.; Gruenberg, H.; Rodewald, K.; Oesterhelt, D.
FEMS Microbiol. Lett. 139, 161-168, 1996
A:Title: A family of halobacterial transducer proteins.
A:Reference number: Z24094; MUID:96275896; PMID:8674984
A:Accession: T46811
A:Status: preliminary;
A:Molecule type: DNA
A:Residues: 1545 <RUD>
A:Cross-references: EMBL:X95589; NID:g1435130; PIDN:CAA64842.1; PID:g1435132
C:Genetics: httpv
A:Gene: httpv
C:Superfamily: Halobacterium salinarum transducer protein htrI
Query Match 8.0%; Score 176.5; DB 2; Length 545;
Best Local Similarity 21.4%; Pred. No. 0.029;
Matches 96; Conservative 79; Mismatches 195; Indels 87; Gaps 17;
QY 12 EYERTPPADLSAQGLEASAAKSAEAGRIAGABAKPKESKTSVERWSILRSVNNALMSL 71
Db 118 EEER-----EAERAREKAEQKQAEER--QTAAEASAKQDARERSABIEQLAADLESQ 169
QY 72 ADKLGASSNSSSTSRGADVSTTATPTPPPT-PPDYKTAQATYDTI-----FTST 125
Db 170 ATEVG-ATLEAASDGLTARVDATTDAENAEIAEVATVNDMLTMTERTIDEIQGFSTNVTI 228
QY 126 SLADIQAALVSLQDVTNKTOT-----KATDEE-----TAAAEWETKNADAVKGAQITE 176
Db 229 ASRETAGAKETQDSQTSVQIEAGTDQDREQLSVAEEMDSYSATVEEVAATAQS 289
QY 177 LAKYASDNQALIDSGLKIT-----SFDLLQAALQSVANNKAAELKEMQDNVPVPGK 230
Db 289 VADTAADTTDVA-TAGQTAEDDAIDAIDAVQETMTQTVANVDALEDLTTEIDD----- 340
QY 231 TPAIAQSLVQDTATATQIEKQGNIRDAYFAGQNASG-----AVENAKNSNSI 279
Db 341 IAEILISDIAEQTNMLA--LNANIEAARAGSGGSGDGFVAVADEVKELATESORSAXDI 398
QY 280 SN-IDSAXAAIATAKTOIAEAKKPPDSPILOAEQMVIAEQLKLNKIPADGSDVPNPG 338
Db 399 AELIEVQSQATTVEEI-----ISQAMDEQAQORSERVSS-----VDDIATISQATAD 445

QY 339 TTVGSGKQGGSIGIRVSMLLDDAENETASIMSGFROMIHMFNTENPDSPQAQQELAA 398
Db 446 ETTDGVQE-----ISQAMDEQAQORSERVSS-----VDDIATISQATAD 484
QY 399 QARAKAGDSDAAALADAQAALGALGKAGQOQGIIN 437
Db 485 RAENVSAASEEQ-TASITEVTSLSQSLAAQADLTLEDRIN 522
RESULT 13
F84194
HtrI4 transducer [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: F84194
R:Ng, M.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl, Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L.
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-627 <STO>
A:Cross-references: GB:AB004437; NID:g10579976; PIDN:AA618922.1; GSPDB:GN00138
C:Genetics:
A:Gene: htrI4
C:Superfamily: Halobacterium salinarum transducer protein htrI
Query Match 7.9%; Score 175; DB 2; Length 627;
Best Local Similarity 22.7%; Pred. No. 0.041;
Matches 116; Conservative 88; Mismatches 190; Indels 116; Gaps 24;
QY 17 PPADLSAQGLEASAAKSAEAGRIAGABAKPK-----ESKTSVERWS-ILRSVNNALMSL 71
Db 145 PALDESVPFAGFESTEMADSLSATLEDTKAELEHQAELEQSQLRALVDALSEA 204
QY 72 ADKLGASSNSSSTSRGADVSTTATPTPPPT-PPDYKTAQATYDTIFTSTSLADIQ 131
Db 205 TD-----AARAGCLTATVDAAALDVTDDHRAAVERDFNQLLETADTISDIQSFSDAV 256
QY 132 AAL-----VSLQDVTNKTOT-----ATDEETAIAAEWETKNADAVKGAQI 174
Db 257 LAVSRTTDERVDAVADRSAAVSESVEIADGANQQTNNLNIAEMDTVSATVEEIAASA 316
QY 175 TELAKYASDNQAIL-----DSLGLK-TSFDLLQAALQSVANNKAAELKEMQDNVPVPG 229
Db 317 NDVAKTA---QAAADRGEDGRGEVEETIEALRALREQQAVAEVETESLAAEVER---IDG 370
QY 230 KTPAIAQSLVQDTATATQIEKQGNIRDAYFAGQNASG-----AVENAKNSNS 278
Db 371 IT-ALIEDIAETNNLAL-----NASIEAARTSGDGFVAVADEVKOLABETREQAAD 423
QY 279 INIDSA-----KKAATATAKTOIAEAKKPPDSP-ILQEAQMVIAEQLKLNKIPKPA-- 329
Db 424 ISEIVDAVTEKAEDASIAIEGYD-AEVEKTKAEGLVLRDFEAIY-----DEVANNHVAQ 478
QY 330 DGSDVPNPG---TTVGSKQGGSSI-----GSIKRVSMLLDDAEN 365
Db 479 EISDATDQGAQSVTVVGMVEVASVSEETAESSTVADNAEQDTATDEVADQMDLAE 538
QY 366 ETASI--LMSGFROMIHMFNTENPDSPQAQQELAAQAAK--AAGDSDSAAALADAQA 421
Db 539 QTPAALAGMLDDF-----TVPADAGTADQSVADDSPTAQPPAADDEPAAAVVDPQPPA 590
QY 422 LEAALGKAGQOQGIINALGQIASAAVVSAG 451
Db 591 SDAE-----DERGVDPDSGGE--SVAVSDGG 613

```

Query Match          7.8%; Score 172.5; DB 2; Length 1128;
Best Local Similarity 23.9%; Pred. No. 0.12;

```

Search completed: March 24, 2004, 06:00:10
Job time : 14.3321 secs

[illegible]

QY 613 PEGSRTPAABIKALSFET 630
Db 1442 -DELATRAKEOKALIAQT 1458

RESULT 8
B90835
probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B90835
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA35073.1; PID:g13361114; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 050952
C:Genetics:
A:Gene: ECs1650

Query Match 7.0%; Score 221.5; DB 2; Length 971;
Best Local Similarity 21.5%; Pred. No. 0.002;
Matches 140; Conservative 98; Mismatches 290; Indels 123; Gaps 20;

QY 26 LEASANKSAEAOIAGAEAKPKESKTSVERWSILRSVAVNLMSLADKLGIASSNSSS 85
Db 110 VEEVARNASAVAQNTAAAKKSADASTSAREATHATDAADSARAASSTAGQAAS-SAQS 168

QY 86 TSSADVSTTATAPPPPTFDYKTAQATDTTFTSLADIQAALVSLQDAVNTIK 145
Db 169 ASSAGTASTKATEASKSAASAAESSKSAATAGAKTSET-----NAAVSQQAAS 223

QY 146 DTAATDEETAIAEWETKNADAVKGAQITELAKYASDNQAILDSLGKLTSPDLQOALL 205
Db 224 TATTKASEAASSA-----RDASAKSAKSETSAASSASAASATAAGNS-----AKAA 274

QY 206 QSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVDQDTATQIEKGNIRDAYFAGON 265
Db 275 KTSETNAKSETAAEQSAAAGSKTAAALSASAASATSA-----GQASASATAAGKS 326

QY 266 ASGAVENAKNNSISNDISAKAATAKTQIAEAKKFPDPSPILOEAEQVIOAEKDLKN 325
Db 327 A-----ESAASSASTATTKAGEATEQ-----ASAAASASAAKTSETN 364

QY 326 IKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDADENETASILMSPROMIHENTE 385
Db 365 AKASSETAESKTAASASASSASAASASSASASK--DEATROASAASKSSATTATKATEAA 422

QY 386 NPDSQAQQELAAQA-----RAKAAGDDSAALADAAQ----- 419
Db 423 GSATAAAQSKTAEASAPRAETAAKRAEDIASAVALDEASTTKKGIQLSSATNSTSESL 482

QY 420 ----KALEAALGKAG-----QQQILNALQIASAAVVSAGVPPA----- 455
Db 483 AATPKAVKAAAYELANGKYTAQDATTAKGIVQLSNATNSTSEMLAATPKSVKAAAYDLANG 542

QY 456 --AASSIGSSVKLYKTSKSGDYKTOISAGIDAYKSIND-AYGR---ARNDATRDVIN 509
Db 543 KYTAQDATTAKGIVQLSSATNSASET-LAATPKAVKAANDNANGRVPSARKVNGKALSS 601

QY 510 NWS-TP---ALTRSVPERATEARG-----PEKTDQALARVIGNSRILGDVYSQVS 556
Db 602 DITLTPKDIGTLNSTMTSGGAGWFKLATVMPQASSVSVITLIGGAGFVNGSP-QQAG 660

QY 557 ALQSWAQIIOQNPOANNEIRKQLTSVATKPPQFGYPYVQLSNDSTOKFIA 607
Db 661 ISELVLRAGNPNKGIITGALWQRTSTGFT-----NPAWNTSGDITYIVA 706

RESULT 9
C85693
probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (stra
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: C85693
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grobbeck, B.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A95480; MUID:21074935; PMID:11206551
A:Accession: C85693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STC>
A:Cross-references: GB:AE005174; NID:g12514847; PIDN:AA656007.1; GSPDB:GN00145; UWGP:Z1
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1918

Query Match 7.0%; Score 221.5; DB 2; Length 973;
Best Local Similarity 21.5%; Pred. No. 0.0021;
Matches 140; Conservative 98; Mismatches 290; Indels 123; Gaps 20;

QY 26 LEASANKSAEAOIAGAEAKPKESKTSVERWSILRSVAVNLMSLADKLGIASSNSSS 85
Db 112 VEEVARNASAVAQNTAAAKKSADASTSAREATHATDAADSARAASSTAGQAAS-SAQS 170

QY 86 TSSADVSTTATAPPPPTFDYKTAQATDTTFTSLADIQAALVSLQDAVNTIK 145
Db 171 ASSAGTASTKATEASKSAASAAESSKSAATAGAKTSET-----NAAVSQQAAS 225

QY 146 DTAATDEETAIAEWETKNADAVKGAQITELAKYASDNQAILDSLGKLTSPDLQOALL 205
Db 226 TATTKASEAASSA-----RDASAKSAKSETSAASSASAASATAAGNS-----AKAA 276

QY 206 QSVANNKKAELLKEMQDNVVPVPGKTPAIAQSLVDQDTATQIEKGNIRDAYFAGON 265
Db 277 KTSETNAKSETAAEQSAAAGSKTAAALSASAASATSA-----GQASASATAAGKS 328

QY 266 ASGAVENAKNNSISNDISAKAATAKTQIAEAKKFPDPSPILOEAEQVIOAEKDLKN 325
Db 329 A-----ESAASSASTATTKAGEATEQ-----ASAAASASAAKTSETN 366

QY 326 IKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDADENETASILMSPROMIHENTE 385
Db 367 AKASSETAESKTAASASASSASAASASSASASK--DEATROASAASKSSATTATKATEAA 424

QY 386 NPDSQAQQELAAQA-----RAKAAGDDSAALADAAQ----- 419
Db 425 GSATAAAQSKTAEASAPRAETAAKRAEDIASAVALDEASTTKKGIQLSSATNSTSESL 484

QY 420 ----KALEAALGKAG-----QQQILNALQIASAAVVSAGVPPA----- 455
Db 485 AATPKAVKAAAYELANGKYTAQDATTAKGIVQLSNATNSTSEMLAATPKSVKAAAYDLANG 544

QY 456 --AASSIGSSVKLYKTSKSGDYKTOISAGIDAYKSIND-AYGR---ARNDATRDVIN 509
Db 545 KYTAQDATTAKGIVQLSSATNSASET-LAATPKAVKAANDNANGRVPSARKVNGKALSS 603

QY 510 NWS-TP---ALTRSVPERATEARG-----PEKTDQALARVIGNSRILGDVYSQVS 556
Db 604 DITLTPKDIGTLNSTMTSGGAGWFKLATVMPQASSVSVITLIGGAGFVNGSP-QQAG 662

QY 557 ALQSWAQIIOQNPOANNEIRKQLTSVATKPPQFGYPYVQLSNDSTOKFIA 607
Db 663 ISELVLRAGNPNKGIITGALWQRTSTGFT-----NPAWNTSGDITYIVA 708

hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: B89921
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizutani-Tsuji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: B89758; MUID:21311952; PMID:11418146
A;Accession: B89921
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-6713 <KUR>
A;Cross-references: GB:BA000018; PID:g13701232; PIDN:BA842527.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: ebha

Query Match 6.9%; Score 220; DB 2; Length 6713;
Best Local Similarity 21.6%; Pred. No. 0.03;
Matches 155; Conservative 106; Mismatches 279; Indels 177; Gaps 32;
QY 27 BASAANKAE-----QRTAGAEAKPKESKTSVERWSILRSVAVNALMSLADKLGIASSNS 82
DB 2200 KASQNSAKSALNGDEKLAAKQTAKSIGRLTLNNAQRTAANAABVDQANLAAVTAAK 2259
QY 83 SSSSTRS-----ADVSTTAT-----APTPPTPTFDYKTAQATYDTTFTSTSL 127
DB 2260 NKATSLNTAMGNLKHAEKNTKESVNYTADQPKQAYDTAVTQAEALINANGSNANE 2319
QY 128 ADIQALVSLQDA-----VTNIKDTAATDEETAIAEWETKNADAVKVGQITELA 178
DB 2320 TQVQAALNQLNAQKNDLNGDKVQAQETA-----KRALASYNLNNAQSTAATSQI----- 2371
QY 179 KYASNQALSLSLGKLTFFDLQALLQSVAN--NNKAELLKEMQDNVVPVPGKTPATAQ 236
DB 2372 ----DNAT-----TVADVTAA-----QNTAELNLTAMQLONGINDQNTV-----K 2408
QY 237 SLVDDOTATATQIEKDGNAIRDAYF--AGONASGAVENAKSNNSISNIDSAAKAAIATAKTQ 295
DB 2409 QQVNTFDADQCK-----KQAYTNVNTAOGILDKANGON-----MTKQAVEAALNQ 2454
QY 296 IAEAKKFPDPILOEAEQWJTQAEKD-----LNKIKPADGSDVPNPGTVGG-- 343
DB 2455 VTTAKNAL-----NGDANVRQAQSDAKANLGLTLHLNNAQKDLTSQIEGATTVNGVN 2507
QY 344 ---SKQOG-----SSIGS---IRVSMLLDDAENETAFILMSGFROMIHMENTE--- 385
DB 2508 SVYTKAQDLDGAVQRLSALANKDQTKASENIDADPTKTKAFDQNAITQAESYLNKDHGT 2567
QY 386 NPDQAQAOQELAAQAPAAKAA--GDDSAALADAKALFAALGKAGQOQGIINALGQIAS 444
DB 2568 NKDKQAVEQAIQSVTSTENALNGDANLQCAKTEATQADNLTLQNTLPQK---TALKQQVN 2624
QY 445 AAVVSGVPPA--AASIGSSVKQLYKTSKTSQSDYKTOISAGY-----DAYKSINDAY 496
DB 2625 AAQRVSGVTDLKNASATSLNNAQDQL-----KQATIGDHDITIVAGGNYTNASPDQKQAYTDAY 2680
QY 497 GRARN--DARTDVNNVS--TPALTRSVPRARTEARG-----PEKTDQALARY----- 540
DB 2681 NAAKINVGSPNVTINADVTAATQTVNNNAETSLNGDTNLTAKQAKALRGWTHLSDA 2740
QY 541 ----ISGNSRTLGDVYSQVSALQSMQIIQS-----NPOANNEIRQKLTSAVTXPK 587
DB 2741 QKQSITQI-----DSATQTVGVQSVKDNA TNLDNMANMQLNSIANKDEVKA----- 2787
QY 588 PQGYFPVQLNSDSTQKF-----IAKLSLFAEGSRTAETIKALSPFNISFIQOVLVN 641
DB 2788 ---SQPYVDADTKQKQAYNTAVTSAENIINATSQPTLDPSAVTQAANQVNTNTALN 2841

RESULT 11

G64887
probable tail fiber protein GP37 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: G64887; T09189
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: G64887
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1122 <BLAT>
A;Cross-references: GB:AE0000234; GB:U00096; NID:g1787633; PIDN:AAC74454.1; PID:g1787633
A;Experimental source: strain K-12, substrain MG1655
R;Aiba, H.; Baba, T.; Fujita, K.; Hayashi, K.; Inada, T.; Isono, K.; Itoh, T.; Kasai, H.; Motomura, K.; Nakade, S.; Nakamura, Y.; Nashimoto, H.; Nishio, Y.; Oshima, T.; Saito, Y.; Horiuchi, T.
DNA Res. 3, 363-377, 1996
A;Title: A 570-Kb DNA sequence of the Escherichia coli K-12 genome corresponding to the
A;Reference number: Z16603; MUID:97251357; PMID:9097039
A;Accession: T09189
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 3-1122 <AIB>
A;Cross-references: GB:AE0000234; GB:U00096; NID:g1787633; PIDN:AAC74454.1; PID:g1787633

Query Match 6.6%; Score 210; DB 2; Length 1122;
Best Local Similarity 20.4%; Pred. No. 0.009;
Matches 154; Conservative 102; Mismatches 294; Indels 206; Gaps 25;
QY 8 GPIDETERTPPA-----DLQAQGLEASAAKSAEAOQTAGAFAKPKESKTSVERWSILRS 63
DB 93 GAMEDDDARPEALRRFFELM---VEEVARNASAVAQNTAAAKSKSDASTSAREANTHAAD 149
QY 64 AVNALMSLADKLGIASSNSSTSRSDADVSTTATATPTPPPTFDYKTAQATYDTTFT 123
DB 150 AADSARAASISAGQAS--SAQASASSAGTASTKATEASKSAAAESSKSAATSAAGAAT 208
QY 124 S--TSLADIQAALVSLQDAVTNIKDTAATDEETAIAEWETKNADAVKVGQITELAKYA 181
DB 209 SETNWSASLQSAATSAATTTKASE--AATSARDAAAKSAEAKSSE-----TNASSSA 259
QY 182 SDNQALDSLGLKTSFDLLQALLQSVANNKKAELLKEMQDNVVPVPGKTPAIQSLVDQ 241
DB 260 SSAASSATAAGN-----SAKAATSETNARSSETATAGQASAAAGSKTAAASASAS 312
QY 242 TDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAAIATAKTQIAEAKQ 301
DB 313 TSA-----GQASASATAAGSA-----ESAASSASTATTKAGEATE 348
QY 302 KFPDSPILQEAQWVIOAEKLNKIPADGSDVPNPGTVGGSKQOQSGISIRVSMLLD 361
DB 349 Q-----ASAAARSASAAKTSETNAKASE-----TSAESSKTAASASSASSAAS 390
QY 362 DAENETASILMSGFPQMTHMENTENPDQAAQELAAQARAAKAGDDSAALADAKA 421
DB 391 ----SAS-----SASAKDEATROASAAKSSATTAATKATEAGSA 427
QY 422 LEAALGKAGQOQGIINALGQIASAAVVSAGVPPAAASSIGSSVKOLYKTSKTSQSDYKTO 481
DB 428 TAAQSKSTAESAATRAETAAKRAEDIASVALEDASTTKKGIVQL---SSATNSTSETL 484
QY 482 ISAGVDYKSIDAYGRARNATRD-----VINNVSTPALT-----RSVPRART 525
DB 485 AATP----KAVKSAVDNAEKRLQKQNGADIPDKGCFLLNNINAVSKTDFADKGRMYRVV 540
QY 526 EA-----RGPEKTDQALARY--ISGNSRTLGD----- 550
DB 541 NAPAGATSKYYPVVMRSAGSVSELASRVIIITATRAGDPMNCFNGFVMPGWTDR 600
QY 551 ---VVSQVSALQSMQIIQSNPQAN--NEEIROKLTSAVTKPPQFGYPVYQLS----- 598

Db 2737 EGASTSASASTSTGASASTSASESASTSASASTSASASTSASASTSASASTSASASTS 2796
QY 541 IGSNSRTLGDVYSQVAGALQVMOQLIQSNQANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
Db 2797 ASESASTSASASTSA--SASASTSASASTSASASTSASASTSASASTSASASTSAS 2841
QY 601 STOKFIKLESLEFAESRTAA 621
Db 2842 STSASVSASTSASASTSAS 2862
RESULT 14
T30822
lmp1 protein - Mycoplasma hominis
C:Species: Mycoplasma hominis
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 07-Dec-1999
C:Accession: T30822
R;Jensen, L.T.; Ladefoged, S.; Birkelund, S.; Christiansen, G.
Infect. Immun. 63, 3336-3347, 1995
A:Title: Selection of Mycoplasma hominis PG21 deletion mutants by cultivation in the pre
A:Reference number: Z18884; MUID:95369882; PMID:7543881
A:Accession: T30822
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1365 <JEN>
A:Cross-references: EMBL:U21962; NID:g790243; PID:g790244; PIDN:AAA81013.1
C:Genetics:
A:Gene: lmp1
A:Genetic code: SGC3

Query Match 6.2%; Score 199; DB 2; Length 1365;
Best Local Similarity 19.1%; Pred. No. 0.04; Mismatches 276; Indels 202; Gaps 28;
Matches 143; Conservative 126;
QY 13 TERPPADLGAQGEASAAKSAQRIAGAEAKPKESKTDVSVERWSILRSVAVNALMSLA 72
Db 317 TDSNKSDESANTELKQALAKANADKV-----QADNLAKSIK 354
QY 73 DKLGIASNSSSSTRSADVSTTATATPTPPPTFDYKTAQATYDTITSTSLADIQA 132
Db 355 EQLNNSVSNANTLSAKLTDKNTIQAQK-----ELEKEVQKA-DQAIKSNNTASMQS 406
QY 133 ALVSLQDAVNIK--PTAATDETAATAAEWETKNADAVKGAQ-----ITELAKY 180
Db 407 AKSLDQKVAEITKKLTFNKKKEAKFNELKQTRNQIQEFINTNKNPNYSELISQTSK 466
QY 181 ASDNQAILDSLGLKTSFDLLQALLQSVANN-----KAAELKEMQDNVPVPGKTPA 233
Db 467 RDSKNSVTDSSNK--SDIESANTELKQALAKANADKVQADNLAKSIKEQLNNSVSNANT-- 523
QY 234 IAQSLVDTQDA---TATQIEKDGNAIRDYAPAGQVAGAVENAKSNNSISNIDSAKAAIA 290
Db 524 LSAKLTDKNTIQAQKTELEKE-----VQKADQAI---KSNNTAS--MQSAKSSL- 568
QY 291 TAKTQIAEAKQKFPDSPILQAEQWVIOAEKDLKNIPADGSDVPNPPTTGGSKQQGSS 350
Db 569 --DAKVAETTKL--ETFNKQKAEKFNELKQTRNQIQEFINTNKNPNY----- 613
QY 351 IGSIRVSMLLDDAENETASILMSGFQMIHMFNTENPDQAAQELAAQAPAAAGDSDS 410
Db 614 --SELISQTSKRDSKN--SVTDSNKSNDIESANTE-----LQALAKANADKVQADNL 663
QY 411 A-----AAALAD-----AQKALEAALGAGQOQILNALQIASAA 446
Db 664 AKSIEQLNNSVSNANTLSAKLTDKNTIQAQKTELEKEIQKANK-----AIKSNNTAS 717
QY 447 VVSAGVPAAPAAASSIGSSVKQLYKTSKSTGSDYKTSAGYDAYKSIINDAYGRAND---- 502
Db 718 MQS-----AKSLDAKVAETTKLTFNKKKEAKFNELKQTRNQIQEFINTNKNPNYS 771
QY 503 -----ATRDVINNVSTPALTRSPVPRARTFARGPEKTDQALARVISGNSRTL-GDYVSQ 554
Db 772 ELISQTSKRDSKNSVTDSSNKSNDIESANTELKQALNTAKAKKSSIDNELPLKNDLQSK 831

QY 555 V-----SALQSV-----MQLIQSNQANNEIR-----QKLTSA 583
Db 832 IBEFGPIRNTNFSWISSKLETTKKNLABELTKADAIKKNPSSKQALXDSQOVOKLQNE 891
QY 584 V-----TKPPQFGYPYVQLSN-----DSTQKFIKLESLEFAE 615
Db 892 LLKTYTEFGKVETKNSNIGYRLFKLQAQEPNNSDVVKLNWEEKQTLSSKKQKLGQ 951
QY 616 GSRTAAEIKALSFETNSLFIQQVLVNI 642
Db 952 STKDYLTQLSTEMSTQESTIKKIVNI 978
RESULT 15
T24583
hypoetical protein T06D8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24583
R;Palmer, S.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z19909
A:Accession: T24583
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1829 <WIL>
A:Cross-references: EMBL:Z49130; PIDN:CAA88964.1; GSPDB:GN00020; CESP:T06D8.1
A:Experimental source: clone T06D8
C:Genetics:
A:Gene: CESP:T06D8.1
A:Map position: 2
A:Introns: 1391/3; 1432/3; 1470/3; 1505/1; 1520/1; 1616/1; 1644/1; 1687/3; 1742/1

Query Match 6.2%; Score 196.5; DB 2; Length 1829;
Best Local Similarity 20.5%; Pred. No. 0.078;
Matches 144; Conservative 108; Mismatches 303; Indels 147; Gaps 26;
QY 18 PADLSA--QGLEASAANKSAQRIAGAEAKPKESKTDVSVERWSILRSVAVNALMSLADKL 75
Db 68 PTELSKDDQVTEAS---GEETTAATAEASSEETTSATVTEGSGEETTVVAVVESGE- 122
QY 76 GIASSNSSSSTRSADVSTTATATPTPPPTFDYKTAQATYDTITSTSLADIQAALV 135
Db 123 EPASSSTSVPTLSKDDQVTEASG-----EETTTAAATEASSEETTSATVTEGSGEE 174
QY 136 SLDQAVNINKDTAAT-----DEETAIAAEWETKNADAVKGAQI 174
Db 175 TTTSAVTEASSEATTTTTPAGTEASGEETTSATVTEGSGEETTVVAVVESGEEPPASSSTSI 234
QY 175 -TELAKYASDNQAILDSLGLKTSFDLLQALLQSVANNKAAELKEMQDNVPV-----P 228
Db 235 PTELSK---NDQVTEASGEET---ITAAATASSETTSATVTEGSGEDTTVVAVVELS 286
QY 229 GKTPALIAQSLVDQTDATQIEKDGNAIRDYAPAGQVAGAVENAKSNNSISNIDSAKAA 288
Db 287 GEQPASSSTSI-----PTELSKD-DQVTEA--SGEETTTAAATEASSEETTSATVTEGSG 337
QY 289 IATAQIABAKKFPDS-----PILQAEQWVIOA-----EKDLKNIP 328
Db 338 EETTVVAVVESGEEPPASSSTSIPTLSKDDQVTEASGEETTTAAATEASSEETTSATVTE 337
QY 329 ADGSDVPNPPTTGGSKQQGSSIGSIRVSMLLDD-----AENETASI 370
Db 398 GSGEDTTVVAVVESGEEPPASSSTSIPTLSKDDQVTEASGEETTTAAATEASSEETTS 457
QY 371 LMSGFQ-----MTHMFNTENPDQAAQ--QELAAQARAKAAGDSDSAAAAALADA--QKA 421
Db 458 VTEGSGEDTTVVAVVESGGEQPASSSTSIPTLSKDDQVTEASGEETTTAAATEASSEETT 517
QY 422 LEAALGAGQOQILNALQIASAAVVSAGVPPAAS--SIGSSVKQLYKTSKSTGSDYKT 480
Db 518 TSAVTEGSGEETTV-----AVVESGEEPPASSSTSIPTLSKDDKDVTEASGEETTT 569

Search completed: March 24, 2004, 06:00:03
Job time : 28.7616 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 19:25:30 ; Search time 17.7616 Seconds
(without alignments)
1892.204 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187
Sequence: 1 MVNPIGPGIDETERTPPAD.....SLFIQQVLVNIQSLYSGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2182	68.5	478	4	US-09-198-452A-776
2	1000	31.4	212	4	US-09-198-452A-775
3	671.5	21.1	361	4	US-09-556-877-299
4	671.5	21.1	361	4	US-09-620-412C-299
5	671.5	21.1	361	4	US-09-598-419-299
6	214.5	6.7	2137	4	US-09-134-001C-4463
7	210	6.6	10182	4	US-09-134-001C-3159
8	186.5	5.9	3696	4	US-09-134-001C-5080
9	178.5	5.6	8991	4	US-08-714-741-32
10	173.5	5.4	1561	3	US-08-894-017-23
11	173.5	5.4	1561	4	US-09-456-474-23
12	173	5.4	2285	4	US-09-308-375-2
13	170.5	5.3	718	4	US-09-540-236-2753
14	167	5.2	1566	2	US-08-687-956A-23
15	165.5	5.2	2310	4	US-08-874-923-120
16	165	5.2	1565	6	5352450-2
17	164.5	5.2	469	4	US-09-489-039A-13565
18	164	5.1	1129	4	US-09-252-991A-29927
19	160.5	5.0	639	1	US-08-466-390-2
20	160.5	5.0	639	1	US-08-470-950-2
21	160.5	5.0	639	1	US-08-467-781-2
22	160.5	5.0	639	1	US-08-195-487-2
23	160.5	5.0	639	1	US-08-483-924-2
24	160.5	5.0	639	5	PCT-US93-06160-2
25	158.5	5.0	468	4	US-09-328-352-6321
26	157.5	4.9	1093	5	PCT-US93-03077-1
27	157.5	4.9	1848	3	US-08-296-791-6

RESULT 1
US-09-198-452A-776
; Sequence 776, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 776
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-776

Query Match 68.5% Score 2182; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 4.5e-144;
Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MVNP	IGPGIDETERTPPADLSAGLEASAAKSAEQAQRIAGAEAKPKESKTDVSVERWSI	60
Db	3	MVNP	IGPGIDETERTPPADLSAGLEASAAKSAEQAQRIAGAEAKPKESKTDVSVERWSI	62
Qy	61	LRSA	VNALMSIADKLGIASSNSSSTSRSDVDSSTTATPTPPPTDDYKTOAQATYDT	120
Db	63	LRSA	VNALMSIADKLGIASSNSSSTSRSDVDSSTTATPTPPPTDDYKTOAQATYDT	122
Qy	121	IFST	SLADIQAALVSLQDAVTNKKDPAETDEETAAAEWETKNADAVKGAQITELAKY	180
Db	123	IFST	SLADIQAALVSLQDAVTNKKDPAETDEETAAAEWETKNADAVKGAQITELAKY	182
Qy	181	ASDN	QAILDSIGKLTSPDLLQALLQSVANNKAAELLKEMQDNFVVPKTPALQAQSLVD	240
Db	183	ASDN	QAILDSIGKLTSPDLLQALLQSVANNKAAELLKEMQDNFVVPKTPALQAQSLVD	242
Qy	241	QTDATATQIEKGN	AIKDAYFAGQNSGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ	300
Db	243	QTDATATQIEKGN	AIKDAYFAGQNSGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ	302
Qy	301	KKFP	DSPIQAEQMVIOAEKDKLNKIPAGSDVNPPTVGGSKQGGSGISIRVSNLL	360
Db	303	KKFP	DSPIQAEQMVIOAEKDKLNKIPAGSDVNPPTVGGSKQGGSGISIRVSNLL	362
Qy	361	DDAENETASII	MSGFRQMIHMFENTENPDSQAQELAAQARAARAKAGDSDAAALADAQK	420
Db	363	DDAENETASII	MSGFRQMIHMFENTENPDSQAQELAAQARAARAKAGDSDAAALADAQK	422


```

QY 421 ALPAALGKAGQQQGIINLALQIASA 447
| | | | | | | | | | | | | | | | | |
Db 423 ALPAALGKAGQQQGIINLALQIASA 449
| | | | | | | | | | | | | | | | | |

RESULT 2
US-09-198-452A-775
; Sequence 775, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 775
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-775

Query Match 31.4%; Score 1000; DB 4; Length 212;
Best Local Similarity 99.5%; Pred. No. 2.9e-62;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 448 VSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKQISAGYDAYKSINDAYGRANDATRDV 507
| | | | | | | | | | | | | | | | | |
Db 9 VSAGVPPAAASSIGSSVNLQYKTSKSTGSDYKQISAGYDAYKSINDAYGRANDATRDV 68
| | | | | | | | | | | | | | | | | |

QY 508 INNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTGLGVYQVSALQSVMLIQS 567
| | | | | | | | | | | | | | | | | |
Db 69 INNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTGLGVYQVSALQSVMLIQS 128
| | | | | | | | | | | | | | | | | |

QY 568 NPQANNEEIROKLTSAVTKPQPGYPVQLSNDSTQKFTAKLSLFAEGSRRTAAEIKALS 627
| | | | | | | | | | | | | | | | | |
Db 129 NPQANNEEIROKLTSAVTKPQPGYPVQLSNDSTQKFTAKLSLFAEGSRRTAAEIKALS 188
| | | | | | | | | | | | | | | | | |

QY 628 FETNSLFIQQLVNLGSLYSGLYQ 651
| | | | | | | | | | | | | | | | | |
Db 189 FETNSLFIQQLVNLGSLYSGLYQ 212
| | | | | | | | | | | | | | | | | |

RESULT 3
US-09-556-877-299
; Sequence 299, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-556-877-299

Query Match 21.1%; Score 671.5; DB 4; Length 361;
Best Local Similarity 42.4%; Pred. No. 4.5e-39;
Matches 154; Conservative 62; Mismatches 130; Indels 17; Gaps 8;

```

```

QY 301 KKFPDPSILQEAQOMVIOAEKDLKNIKPADGSDVPNPGTTVGSKQQGSSIGSI----- 354
| | | | | | | | | | | | | | | | | |
Db 2 QEIADSPLVKAAEQINOAQDIIQTITPS-GLDIPVGPS-GSAASAGSAGALKSSNNS 59
| | | | | | | | | | | | | | | | | |

QY 355 -RVSMILLDDAENETASILMSGFROHMENTENPDSQAQQELAAQARA---AAGDGS 410
| | | | | | | | | | | | | | | | | |
Db 60 GRISLLDDVDNEMAAIAAQGFPSMIEQFNVNPNATAKELQAMEAQLTAMSDQLVGADGE 119
| | | | | | | | | | | | | | | | | |

QY 411 AAAALADAQKALEAALGKAGQQQGIINLALQIASA 449
| | | | | | | | | | | | | | | | | |
Db 120 LPAEIOAIKDALAQAAL-KOPSADGLATANGQVAFAAKVGGSAGTAGTIVQMNVLKLYKT 178
| | | | | | | | | | | | | | | | | |

QY 471 --SKTSGDYKTIQISAGYDAYKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTAR 528
| | | | | | | | | | | | | | | | | |
Db 179 AFSSTSSSSYAAALSDDGYSAYKTLNLSYSEKRS-GVQSAISQTNPALSRVSRSRGIESQ 237
| | | | | | | | | | | | | | | | | |

QY 529 G-PEKTDQALARVISGNSRTGLGVYQVSALQSVMLIQS 627
| | | | | | | | | | | | | | | | | |
Db 238 GRSADASQRAAEITVRDSTGLGVYQVSALQSVMLIQS 297
| | | | | | | | | | | | | | | | | |

QY 588 PQFGYPVQLSNDSTQKFTAKLSLFAEGSRRTAAEIKALS 647
| | | | | | | | | | | | | | | | | |
Db 298 PQFGYPVQLSNDSTQKFTAKLSLFAEGSRRTAAEIKALS 357
| | | | | | | | | | | | | | | | | |

QY 648 GYL 650
| | | | | | | | | | | | | | | | | |
Db 358 GYL 360
| | | | | | | | | | | | | | | | | |

RESULT 4
US-09-620-412C-299
; Sequence 299, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-299

Query Match 21.1%; Score 671.5; DB 4; Length 361;
Best Local Similarity 42.4%; Pred. No. 4.5e-39;
Matches 154; Conservative 62; Mismatches 130; Indels 17; Gaps 8;

```

Db 298 PCGYPAVQNSVDSIQKFAAQIEREFVDGERSLAESQENAFRKQPAPITQQVLNIA SLFS 357
648 GYL 650
Db 358 GYL 360

```

RESULT 5
US-09-598-419-299
; Sequence 299, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Scholler, John
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598,419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-598-419-299

```

RESULT 6
 US-09-134-001C-4463
 ; Sequence 4463, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ;
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS AUREUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964

[illegible]

RESULT 6
US-09-134-001C-4463
; Sequence 4463, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964

FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5080
LENGTH: 3696
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match 6.6%; Score 210; DB 4; Length 10182;
Best Local Similarity 20.4%; Pred. No. 6.4e-05;
Matches 149; Conservative 108; Mismatches 290; Indels 184; Gaps 30;
QY 16 TTPADLSAQGLESAANKSAEAGRI-----KPKES-----KTDV 55
Db 2628 TPQVNSALSQVQAQKINEAKALLQNKADNSQLVRAKEQLQSQIQAASDTGMDTQDST 2687
QY 56 ERWSILRSVAVNLSADKLGIASSNSSTSRSDVSTTATPTPPPTFDY-----110
Db 2688 RNYKVRQAQAEQIAHANSV---INNGATSQQINDAKNIVEQAQR-----DYVEAKS 2737
QY 111 -----KTQAQATYDTIFTSTSLADIAQALV-SLQDAVNTNIKDTAATDEETAIAAEWETKN 164
Db 2738 NLRADKSQLQSAVDLNRDLTNDKPKASVRRYNEAISNR-----KE 2780
QY 165 ADAVKVGAQITELAKIYASDQALDLSGLKLTSPD--LLQA-ALLOSVAANNKAAELLKEM 221
Db 2781 LDTAKADASST-LRNTNPSVEQVRDALNKINTQPKVQNALLOPKENNELVQAQKRL 2839
QY 222 QDN-----PVVPGKTP-----ALAQSLVQDTATATQIEKGNNAIRDAY 260
Db 2840 QDAVNDIPQTCGTQQTQNTNNYNDKQREARALTSQRVDNGDATTCETISEKSKVEQAM 2999
QY 261 FAGONASG-----AVENAKNN-----SISNIDSAKAAIAKTAQIABAQK 301
Db 2900 QALTNKASNLRAKDELQATYKLIENVTNGKPKASIROVETAKARI---QNDK 2956
QY 302 KFPDSPILQAEQVQIAEADKLNKIPADGSDVNPCTTGGSKQCGSSIGSIRVSM 361
Db 2957 E-AERILGNDNPQVSVQTAQNKIKAIQ-----PKLTEINMLQKNNTELVAKNRLE 3010
QY 362 DAENETASILMSGPROMIHMENTNPDSQAQQLAAQARAQAAAGDSDSAAALADAQA 421
Db 3011 NAVNDTDP--THGWTQ--ETINNNYAKREAQNEI--QKANMLNNGDATAQDTSSEKSK 3064
QY 422 LEAALGKAGQOQGLNALGQIASAAVSVAGVPPAAASSIGSVKQLYKTSKSTG-----475
Db 3065 VEQVL-----QALQNAKNDLR-----ADKRELOTAYNKLIQNVNTNGKPKSSI 3107
QY 476 SDYKTIQISAGYDAVKSINDAYGRARND-----TRDVIN-----NVSTELTRSV 520
Db 3108 QNYKS-----ARRNIENQYNTAKNAHNVLENTNPTVNAVEDALBKINAIQEVTKAI 3160
QY 521 PRAKTE-----ARGPKTDQALARVIGNSRTIGDVYSQVSAQSVMLIQS-----N 568
Db 3161 NILQDKEDNSLVRKEKLDQAINSQPSLNGMTQESINNYTTKREQAQNTASSADTIINN 3220
QY 569 PQANNEETR-----QKLTSAVTTPPQPGYGYVVLNSDSTOKFTAKLESFAEGSRFAE 622
Db 3221 GDASIEQTEINKIRVEEATNALNEAKQ-----HLTADTTSLKTEVRKLRRRGDTNNKPKSS 3276
QY 623 IKALSFFETNSL 633
Db 3277 VSAYNNTIHS 3287

RESULT 8
US-09-134-001C-5080
Sequence 5080, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5080
LENGTH: 3696
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5080

Query Match 5.9%; Score 186.5; DB 4; Length 3696;
Best Local Similarity 21.7%; Pred. No. 0.00067;
Matches 171; Conservative 118; Mismatches 293; Indels 205; Gaps 40;
QY 12 ETERTPPADLSAQGLESAANKSAEAGRI-----AGAEAKPKESKTDVSVERNSILR----62
Db 1956 QINNTTDA-TEEEKQEAATNRVNAQLAQATQINNAHSTOEVAESKSTNATTKSVQPNVI 2014
QY 63 ---SAYNALMSLADKLGIASSNSSTSRSDVSTTATA-----PTPPPTFD 109
Db 2015 KPETAINSLTQANNOKTLGNDGNATDDKEAKAKLVTKLNEQIQKIHESQODNOVDN 2074
QY 110 YKTAQATA-----YDTIFTSTSLA-----DIOAALVSLQDAVNTNIKDTA--ATD 151
Db 2075 VKQAATAIKLINANAHRQDAINILNLAESKSDIRAN---QDATTEKNTAQSID 2130
QY 152 EETA-----IAEWETK---NAD-AVKVGAQITELAK 179
Db 2131 DTLQAQNNINGANTVALVDENLEDQKQLRVLSTQTKQAKADIQAIGQORSTIDQ 2190
QY 180 ---VASDQAIILDSGLKTS---FDLLQALL--QSVAN--NNKAAELLKMDQNPVV---P 228
Db 2191 NQWATTEEQEALERLNQETNGVNDRIQAALANQVNTDEKNWILETIRNVE--PIVIVKP 2248
QY 229 GKTPTAQSLVQDTATATQIEKGNNAIRDAYPAGNASGAVENAKSN--NSISNTDSAK 286
Db 2249 KANEIIRKAAEQT---TLNQNDATLE---EKQIALGKLEVEKNEALNQVSAHSNN 2301
QY 287 AATATAKTOIAEAQKFPDSPILQAEQVQI-----AEKDLKN 325
Db 2302 DVKIAENNGIAKISEVHPETIKRKAQIEBQDAQSQIDITINANNKSTNEEKSAAIDRVN 2361
QY 326 IKPADG--SDVNPNGTT--VGGSKQCG--SSIGSI-----RVSMLLDDAENETASILM 372
Db 2362 VAKIDAINNTTATTTQLVNDKXNSGNTSISQILPSTAVKTNALAALAEAKKNKNAII--2419
QY 373 SGFRQMIHMENTNPDSQAQQLA-----QAARAQAAGDSDSAAALADAQKAL 422
Db 2420 ---DQTPNATAEEKEANKNVDRLOQEEADANILKAHTTDEVNNIKNQAVNI 2468
QY 423 EAALGKAGQOQGLNALGQ-IASAAVSVAGVPPAAASSIGSVKQLYKTSKSTGSDYKTQ 481
Db 2469 NAVQVEVIKQNVKQNLQFIDNQKLIENTPDTLEEKAEANRLQLNVLTSDSEI---2525
QY 482 ISAGYDAYKSINDAYGRARN-----DATKDVNNVSTP--ALTRSVPRARTBARGP 530
Db 2526 --ANVDHNEVDQALDKARPKIEEIVPQVSKKRDVLNAIQEAQFNSQTEI-QENQEA 2582
QY 531 EKTQ-----QAL--ARVIGNSRTIGDVYS-QVSALQSVMLIQSNPQAN-----NE 574
Db 2583 EKTEALNKINQLNQAKVNDQASNKDVSATKRSIQDIEQ-IQHPQTKATGRHLNE 2641
QY 575 EIRQKLTSAVTKPPQPGYGYVVLNSDSTOKFTAKLESFAEGSRFAEIKALSFFETNSLF 634
Db 2642 KANQOQSTIATHP-----NSTIEERQESAKLQEVLLK-----AIKIDKQIND-D 2687
QY 635 IQOVLVN 641
Db 2688 VEKTVVN 2694


```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-894-017-23

Query Match      5.4%; Score 173.5; DB 3; Length 1561;
Best Local Similarity 19.9%; Pred. No. 0.0016;
Matches 140; Conservative 118; Mismatches 304; Indels 143; Gaps 27;

Qy 48 KESKTSVERMSILRSVAVNALMSLADKLGIASSN-SSSSTSRSDVSTTATPTPPPT 106
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 KVKKTYGFRKSKISKTLGAVLGTVAAVSVAGQVFADETTTSDVTKVVGTCGNPAT 61
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
107 -----FDYKTKQAQYDTITFTST-----SLADIQAALVS-LQDAVNIKDT 147
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 NLPEAQSASQAQESQTKLERQMVHTIEVPKTDLDQAADKASGAVNVVQDADVN-KGT 120
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 AATDEETAIAAEWE-----TKNADAVK-----VGAQITELAKYASDNOAILDSLKG 193
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 VKTAE-E-AVOKETEIKEDYTKQAEIKKTTDQYKSDVAHEAEVAKIKAKNQATKEQYK 179
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 LTSFDLL-QAALLQSVANNKAAEL-----LKEMQDNVPVPGKTPAIAQSLVDQT----D 243
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 ----DMVAHKAEVERINAANAASKTAYEAKLAQYQADLAQVKTNAANQASQKALAAQ 235
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
244 ATATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQKFP 303
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
236 AELKRVQEAANAQAAY----DTAVAANNAKNTETIAAANEIEIRKNATAKA-----282
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
304 PDSPILOEAEQMWIOAEKDLKNIKPADGSDVPNPPTTVGG-----SKQGSISGIR 355
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
283 -----EYETKLAQYQAEKRVQEAANAANEADYQAKLTAYQTELARVQKANADAKAAYE 335
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
356 VSMILLDOENETASILMSGFQTHMNTENPDQAAQOQELAAQARAAGAAGDSDSAAAL 415
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
336 AAVAANNAKNAALTAENTAIAKQ-----RNEAKATYEAAALKQYEADLAQVKKANAANE 388
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
416 ADAQKALEALGKAGQOQILNALGQIA-SAAVVSAGVPPAAASSIGSSVKQ-----466
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
389 ADYQAKLTAYQTELARVQK-ANADAKAAYEAAVAANNAANALTAENTAIAKKNADAKAD 447
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
467 -----LYKTSKSTGSDYKTOISAGYDAYKSIDAYGRANDATRDVINNVSTPA 515
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
448 YEAKLAKYQADLAKYQDL-ADYPVKLKAYEDEQASIKAAALEKHKNE--GNLTPTS 504
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
516 LTRSVPRARTARGPEKTDQALARVISGNSRTLGDVYSQVSL-----QSVMQ-----563
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
505 AQNLVYDLEPNANLSLTDDGKFLK-----ASAVDDAFSKSTSKAKYDQKILQLDLDTN 559
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
564 IIQSNPOANNEEI-----RQKLTSAVTKPPQFGYPYVQLS-NDSTQKFIKLSLFAEG 616
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
560 LEQNDVASSMELYCNFGDKAGWSTTVSNNSQVWKGSVLLERQGSATATYTNLQNSYNG 619
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
617 SRTAAEIKALSFETNS-----LFIQQLVNLVIGSLYSGYLQ 651
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
620 KKISKIVKYTVDPKSKFQGWLGIFDPTDLGVFASAYTCQVE 664

RESULT 11
US-09-456-474-23
; Sequence 23, Application US/09456474
; Patent No. 6500433
; GENERAL INFORMATION:
; APPLICANT: Lehner, Thomas
; APPLICANT: Kelly, Charles
; TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE OF
; COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II
; FILE REFERENCE: 25150-20067-10
; CURRENT APPLICATION NUMBER: US/09/456,474
; CURRENT FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: US 08/894,017
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 23
; LENGTH: 1561
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-456-474-23

Query Match      5.4%; Score 173.5; DB 4; Length 1561;
Best Local Similarity 19.9%; Pred. No. 0.0016;
Matches 140; Conservative 118; Mismatches 304; Indels 143; Gaps 27;

Qy 48 KESKTSVERMSILRSVAVNALMSLADKLGIASSN-SSSSTSRSDVSTTATPTPPPT 106
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2 KVKKTYGFRKSKISKTLGAVLGTVAAVSVAGQVFADETTTSDVTKVVGTCGNPAT 61
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
107 -----FDYKTKQAQYDTITFTST-----SLADIQAALVS-LQDAVNIKDT 147
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 NLPEAQSASQAQESQTKLERQMVHTIEVPKTDLDQAADKASGAVNVVQDADVN-KGT 120
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 AATDEETAIAAEWE-----TKNADAVK-----VGAQITELAKYASDNOAILDSLKG 193
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 VKTAE-E-AVOKETEIKEDYTKQAEIKKTTDQYKSDVAHEAEVAKIKAKNQATKEQYK 179
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
194 LTSFDLL-QAALLQSVANNKAAEL-----LKEMQDNVPVPGKTPAIAQSLVDQT----D 243
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
180 ----DMVAHKAEVERINAANAASKTAYEAKLAQYQADLAQVKTNAANQASQKALAAQ 235
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
244 ATATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAQKFP 303
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
236 AELKRVQEAANAQAAY----DTAVAANNAKNTETIAAANEIEIRKNATAKA-----282
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
304 PDSPILOEAEQMWIOAEKDLKNIKPADGSDVPNPPTTVGG-----SKQGSISGIR 355
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
283 -----EYETKLAQYQAEKRVQEAANAANEADYQAKLTAYQTELARVQKANADAKAAYE 335
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
356 VSMILLDOENETASILMSGFQTHMNTENPDQAAQOQELAAQARAAGAAGDSDSAAAL 415
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
336 AAVAANNAKNAALTAENTAIAKQ-----RNEAKATYEAAALKQYEADLAQVKKANAANE 388
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
416 ADAQKALEALGKAGQOQILNALGQIA-SAAVVSAGVPPAAASSIGSSVKQ-----466
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
389 ADYQAKLTAYQTELARVQK-ANADAKAAYEAAVAANNAANALTAENTAIAKKNADAKAD 447
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
467 -----LYKTSKSTGSDYKTOISAGYDAYKSIDAYGRANDATRDVINNVSTPA 515
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
448 YEAKLAKYQADLAKYQDL-ADYPVKLKAYEDEQASIKAAALEKHKNE--GNLTPTS 504
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
516 LTRSVPRARTARGPEKTDQALARVISGNSRTLGDVYSQVSL-----QSVMQ-----563
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
505 AQNLVYDLEPNANLSLTDDGKFLK-----ASAVDDAFSKSTSKAKYDQKILQLDLDTN 559
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
564 IIQSNPOANNEEI-----RQKLTSAVTKPPQFGYPYVQLS-NDSTQKFIKLSLFAEG 616
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
560 LEQNDVASSMELYCNFGDKAGWSTTVSNNSQVWKGSVLLERQGSATATYTNLQNSYNG 619
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
617 SRTAAEIKALSFETNS-----LFIQQLVNLVIGSLYSGYLQ 651
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
620 KKISKIVKYTVDPKSKFQGWLGIFDPTDLGVFASAYTCQVE 664

RESULT 12
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394-PCT
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-308-375-2

Query Match      5.4%; Score 173; DB 4; Length 2285;
Best Local Similarity 20.4%; Pred. No. 0.003;
Matches 171; Conservative 126; Mismatches 292; Indels 250; Gaps 37;

QY 20 DLSAOLE--ASAANKSAEAOIA--GAEAKPKESKTD---SVERWSILRSVAVNALMSL 71
Db 11 DDPLELEKFSLSGLKQEMQKALDSGDEKAFDPAKQLOSLLETYSKSDSIDVFKMS 1040
QY 72 ADKL--GIASSNSSSTSR--ADVSTTATPTPPPTFPDDYKTAQPAYDTIETSTSL 127
Db 1041 FDKAQNKIKDQKSLSSVKSEVSGDGETLAEAGNEA---EDFGKLEKALD---ANSV 1092
QY 128 ADIQAALVSLQAV---TNIKD-----TAATDEETAI 156
Db 1093 DDIAKAEMSDAMQFDSVDVLNGDFNNTKDQVAPLNDLLEKMAEGKSIANEANTLI 1152
QY 157 AAEMETKNADAVKGA-----QITELAKYASD-----NOAI----- 187
Db 1153 QKDKELAQAISLGVVVKINRDEVIKQKVKLDAYNDVTVYSNKLMTKTEVNAIKTLNAD 1212
QY 188 ---LDSLGLK---TSPDLQAAL---LOSVAANNKAABELKEMQNPVVPVPG----- 229
Db 1213 TLRIIDSLKKLRKERKLDMSAEISDLEVKSIINNADAKKELKLEKMLQPGGYSQIE 1272
QY 230 ----KTPATAQSLVDOTATATQIEKDGNAIRDAYPAGQNASGAVE---NAKSNNSISNTD 283
Db 1273 AMQSVKSALESYISABEATSTQ-EMNKQALVEAGTSLNWTDDQKANEETKTSMYVVD 1331
QY 284 SAKAAATATAKTOIAEAKPKFPDPS-----ILOAEQMWIOAEKDLKN--- 325
Db 1332 KYKEALEKYNVAEIDKYNQVNDYPKYSQYKRDAIKEIKALQOKKMLQOEAKLLKQOLK 1391
QY 326 -----TKPADGSDVPNGPTVGG-SKQGGSSIGSIRVSMILLDAENETASILMSGFR 376
Db 1392 SGNITQYGIPTSTSSGGTSPSTGGSYSGKYSYINSAASKYNVDPD-----LIAAVI 1444
QY 377 QMTHMENTENPDSQAQOEALAAQARAARAKAGDDSAALADAQ---KALEAALGK-AGQQ 432
Db 1445 QQESGENAKARSGVGAWGLMQLMPATAKSLGVNNAYDPYONVWGGTKYLAQOLEKFGNV 1504
QY 433 QGITNALGQIASAAVVSAGVPP-----AAASIGSVKQLY----- 468
Db 1505 EKALAAYNAGPGNVIKYGGIPPPKETQNYVKIMANYKSLSSATSSIASYYTNNSAFRV 1564
QY 469 -----KTSKSTGSDY-----KTQISAGVDAYKSIDAYGRARND 503
Db 1565 SSKYGOQESLRSRPHKGTDFAAKAGTAKTSLQSGKVQI-AGYS--KTAGNWWVIKQDDG 1621
QY 504 T-RDVNNVSTPALT-RSVPRART-----EARG-----PEKTDQAL 537
Db 1622 TVAKYXHMNTSPKAGQSVKAGQTTGKVGSTGNSTGNHLHLQIEQNGKTIPEKTMQGI 1681
QY 538 ARVISGNSRTLGIVYQVSGALQVWMOIQSNPQANNEIRQKLTSAVTKPPQGYPVQL 597
Db 1682 GTSISDASQAERQOGIAQAKSDLLSLQGDISSVNDIQ-----ELQVELVQS 1730
QY 598 SINDSTOK-----FTAKLESB---FAEGR-----TAARIKALSRETSLFTQQVLVN 641
Db 1731 KLDEFDKRIGDFDVRIAKDESMANRYTSDSKERPKYTSQCKKAVAEQAK---IQQOKVN 1786
```

RESULT 13

```
US-09-540-236-2753
; Sequence 2753, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
```

RESULT 14

```
US-08-687-956A-23
; Sequence 23, Application US/08687956A
; Patent No. 5861157
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAF
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2753
; LENGTH: 718
; TYPE: PRT
; ORGANISM: M.cattarrhalis
US-09-540-236-2753
```

```
Query Match      5.3%; Score 170.5; DB 4; Length 718;
Best Local Similarity 19.2%; Pred. No. 0.0089;
Matches 144; Conservative 120; Mismatches 248; Indels 237; Gaps 36;

QY 11 DETERTPPADLSAQL-----EASAANKSAEAGR-----IAGAEAKPKESKTD 53
Db 32 DDVEIVAEVDIDNQAQVSVLIIRDITDKADQADHTDDASKADDETIVDGVKQKQAEKED 91
QY 54 SYVERWSILRSVAVNALMSLADKLGIIASSNSSSTSRSDVSTTATATPTPPPTFPDDYKQ 113
Db 92 FENKAQDLQD-----KATEKLEVAKEATQDKVEKTQSL-----VEDIKDK 131
QY 114 AQATYDTTFTSLADIQAALVSLQDAYTNIKDTAATDEETALAFAEWETKNADAVKGAQ 173
Db 132 AQSLQE-----DAADTVLEALKQ-AASDKVETTKAEQSLKDDA---TQ 170
QY 174 ITELAKYASD-----NOAILDLSGLK-----TSPD-----LIQAALLQSVANNKA 214
Db 171 TPESAKQAVEGKVEAIKQVLDQVDSLKDDTDQDNTDQOEKQTLKDXAVQAATAAKRV 230
QY 215 AELLKEMQNP-----VVPKTPALIAQSLVDTATATQIEKDGNAIRDAYPAGQNASGA 269
Db 231 EDVDDVXHTTSPKNTASGKIDEIKQAADVKTBEVKSQSKADALKS---SGBELKQT 287
QY 270 VENAKSNNSISNIDSAKALATAKTQIAFAOKKFPDPSILOAEQMWIOAEKDLKNI--- 326
Db 288 AGTA-ANDAIT---EAQAAVVGSGVAAA-----DSAQSTQASAKDKLNQLE 330
QY 327 --KPADGSDVPNGPTVGGSKQGGSSIGSIRVSMILLDAENETASILMSGFRQIMHMENT 384
Db 331 QGKSALDEKVOELGEKFGATE-----KINAVSENVDLATQVI----- 367
QY 385 ENPDQAAQOEALAAQARAARAKAGDDSAALADAQALEAALGKAG----- 430
Db 368 -KEEAQALQTNAGESLQAAKAAGEBYDA---THEDKGLTTLKLGKVGAVLSGMYGISONKN 423
QY 431 -QOQGI-----LNA-----LGOIASAAVVSAG-----VPPAAASIGSSVKOL 467
Db 424 KHYQGVDLHRESFDKDAFHAQSSFFAGQIFGAKAVAKNVAAKVFPQSKFEAIGES---L 480
QY 468 YKTSKSTGSDYKTOISAGVDAYKSIDAYGRARNDATRDVINNVSTPALTTRSVPRARTEA 527
Db 481 Y-----NKVAENSNAM-AIKD---LKNDRPDLINTMNT-----QERHAPA 517
QY 528 RGEPTKDOALRV-----ISGNSRTLGD-VYSQVSGALQSVMOI---IQSNPQANNEIRQK 579
Db 518 EDVANQNRLATLGGVAGLAGLKGVLADAAMLLMVLSTRTVYQVAAIYDQPLTGKEGKTK- 576
QY 580 LHSVTKPPQFGYPYVQJLNDSTQKF-----LAKLESIFAEGSRTA--AEIKALS- 627
Db 577 -----AYGVLSGANLEKQKQVILTALGSSMLANAQQTGIKAQLDLSLA 623
QY 628 -PETSLSFTQQVL-----VNIGLSYGYL 650
Db 624 RYRESQPYAKQFLDLDKFNLDNLNPNML 652
```

RESULT 14

```
US-08-687-956A-23
; Sequence 23, Application US/08687956A
; Patent No. 5861157
```

```

GENERAL INFORMATION:
APPLICANT: BURNIE, JAMES P
APPLICANT: MATTHEWS, RUTH C
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF
TITLE OF INVENTION: INFECTIONS DUE TO STREPTOCOCCI AND ENTEROCOCCI
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, LLP
STREET: 1100 NEW YORK AVENUE, N.W., EAST TOWER, 9TH
FLOOR
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,956A
FILING DATE: 29-JUL-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9401689.6
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, PAUL N
REGISTRATION NUMBER: 16,773
REFERENCE/DOCKET NUMBER: 50885/222892
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/861-3000
TELEFAX: 202/822-0944
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1566 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Streptococcus sobrinus
US-08-687-956A-23

Query Match 5.2%; Score 167; DB 2; Length 1566;
Best Local Similarity 21.7%; Pred. No. 0.0046;
Matches 131; Conservative 91; Mismatches 245; Indels 136; Gaps 22;

Qy 48 KESKTSVERWSILRSVAVNALMSLADKLGIASSNSSSTS--RSADVSTTATPTPTPP 105
Db 10 RKSISRITLQALLGTAI--LASVTGQKALAEETSTTSGVNTAVVGTETGNPATNLPD 67
Qy 106 TFDYKTKTAQTAY-----DTFTSTSLADIQAALVSLQDAVINIKDTATD----- 151
Db 68 KQNPSSQAETSQAQCKTGAMSDVSTSELDEAKSAQBAQVTSQDATVDKGVETS 127
Qy 152 -----ETATAAEWETKNADAVKGAQITELAKYA--SDNOAILDSLGLTSFLLQALL 205
Db 128 DEANQKETEIKDDYSKQAADIKQT-----TEDIKAARVNRQAETD---RITQENAKKAQY 180
Qy 206 QSVANNKAELLKEMQNPVVPKTPAIAQSLVDQTDATATQIEKQGNAIRDAYFAGQN 265
Db 181 EQDLAANK-ABVERITNEN-----AQKADY-EAKLAQYQKDLAAVQQAANDSQA 228
Qy 266 ASGAVENAKSNNSISNIDSAKAATATAKTQIAEA-----QKKEPDSPILO----- 310
Db 229 AYAAAKEA-YDKELARVQAANAA--AKKEYEALAAANTTKEQIKAEANAIQQRNAQK 284
Qy 311 -EAEQWITQAEKOLKLNKIPADGSDVPNGTTVGSGKQGGSSIGSVIRVSMILLDDAENETAS 369

```

```

Db 285 ADYEAKLAQYKDL-----AAAQSG-----NATNE----- 309
Qy 370 ILMGFRQMIHMFNTENTPDQAQOELAAQARAAXAAGDDSAALAA--DAQKALEAALG 427
Db 310 -----ADYQAKKAAYQELARVQAANAAKQAVEQALAAANTAKNAQITAEN 355
Qy 428 KAGQQGIIINALGQITASAAVVSAGVPPPAASISGSSVKOLYKTSKSTGSDYKTQISAGYD 487
Db 356 EAIQQR--NAQAKANYEAKLAQYQKDLAAQSGNAANE-----ADYQEKLAAYEK 403
Qy 488 AVKSNIDAYGRANDATRDVINNVSPALTRVSPARTAEARGPEKTDQALARVNSGRT 547
Db 404 ELARVQAANAAKQVEYQVQEAANKAGNBEITANRAIRERNAKAKTDYELK--LSKYQEE 461
Qy 548 LGDVYSQVSALQSVMOI IQSNFPANNEEIRQKLTSAVTKPPQPGYPYVOLSNDSQKFTA 607
Db 462 LAQYKDLAEYPAKLQAYQEQAA-----IKAALLELEKHKHEDWNLSPPSAQSLVY 513
Qy 608 KLE 610
Db 514 DLE 516

RESULT 15
US-09-874-923-120
Sequence 120, Application US/09874923
Patent No. 6638517
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Webb, John R.
APPLICANT: Dillon, Davin C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Bhatia, Ajay
APPLICANT: Coler, Rhea
APPLICANT: Probst, Peter
APPLICANT: Brannon, Mark
TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
FILE REFERENCE: 210121.420C8
CURRENT APPLICATION NUMBER: US/09/874,923
CURRENT FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 120
LENGTH: 2310
TYPE: PRT
ORGANISM: Leishmania major and chagasi
US-09-874-923-120

Query Match 5.2%; Score 165.5; DB 4; Length 2310;
Best Local Similarity 19.9%; Pred. No. 0.01;
Matches 134; Conservative 122; Mismatches 295; Indels 121; Gaps 24;

Qy 12 ETERTPPADLSAAGLEASAANKSAEQAQRIAGAEKPKESKTSVERWSILRSVAVNALMSL 71
Db 984 DTATQORAELEAQLARLAADREARQQLAANAEELQORLDTATQOR-AEAEQAVRLAAN 1042
Qy 72 ADKLGIASSNSSSTSRSADVSTTATPTPTPTPTFDYKTKTAQTAYDTFTST-SLADI 130
Db 1043 ABEL---QORLDTATQORAELEARVARLAADREARQQLAANAEELQORLDTATQORAE 1099
Qy 131 QAALVSLQDAVINIKDTATATDEBETAATAEWETKNADAVKGAQITELAKYASDNQAILDS 190
Db 1100 EARVABL-----AANAEE--LQORLDTATQORAELEAQLARLAANAEELQORLDT 1147
Qy 191 LGLKTSFDLLQAALQSVANNKKAELL---KEMQNPVVPKTPAIAQSLVDQTDATA 246
Db 1148 ATQORA--ELEEARVARLAADREARQQLAANAEELQORLDTATQORAELEAQLARLAANA 1205
Qy 247 TQIEK-DGNAIRDAYFAGQNAAGAVENAKSNNSI--SNIDSAKAATATAKTQIAEAKQKF 303
Db 1206 EELQORLDTATQORAELEAQLARLAADREARQQLAANAEELQORLDTATQORAELE--- 1262

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 24, 2004, 05:59:21 ; Search time 152.132 Seconds
(without alignments)
1108.119 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187
Sequence: 1 MWNPIGPGPIDETERTPPAD.....SLFIQVLVNIISLYSGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpa/US07_PUBCOMB.pap.*
 - 2: /cgn2_6/ptodata/2/pubpa/PCT_NEW_PUB.pap.*
 - 3: /cgn2_6/ptodata/2/pubpa/US06_NEW_PUB.pap.*
 - 4: /cgn2_6/ptodata/2/pubpa/US06_PUBCOMB.pap.*
 - 5: /cgn2_6/ptodata/2/pubpa/US07_NEW_PUB.pap.*
 - 6: /cgn2_6/ptodata/2/pubpa/PCTUS_PUBCOMB.pap.*
 - 7: /cgn2_6/ptodata/2/pubpa/US08_NEW_PUB.pap.*
 - 8: /cgn2_6/ptodata/2/pubpa/US08_PUBCOMB.pap.*
 - 9: /cgn2_6/ptodata/2/pubpa/US09A_PUBCOMB.pap.*
 - 10: /cgn2_6/ptodata/2/pubpa/US09B_PUBCOMB.pap.*
 - 11: /cgn2_6/ptodata/2/pubpa/US09C_PUBCOMB.pap.*
 - 12: /cgn2_6/ptodata/2/pubpa/US09_NEW_PUB.pap.*
 - 13: /cgn2_6/ptodata/2/pubpa/US10A_PUBCOMB.pap.*
 - 14: /cgn2_6/ptodata/2/pubpa/US10B_PUBCOMB.pap.*
 - 15: /cgn2_6/ptodata/2/pubpa/US10C_PUBCOMB.pap.*
 - 16: /cgn2_6/ptodata/2/pubpa/US10_NEW_PUB.pap.*
 - 17: /cgn2_6/ptodata/2/pubpa/US60_NEW_PUB.pap.*
 - 18: /cgn2_6/ptodata/2/pubpa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	9	US-09-841-132-405
2	3187	100.0	651	10	US-09-564-479-2
3	3187	100.0	651	10	US-09-841-260-78
4	3187	100.0	651	13	US-10-007-693-78
5	3187	100.0	651	15	US-10-312-273-13
6	2844	89.2	583	10	US-09-564-479-4
7	2208	69.3	490	10	US-09-564-479-8
8	2204	69.2	452	10	US-09-564-479-6
9	2182	68.5	478	15	US-10-289-762-776
10	1000	31.4	212	15	US-10-289-762-775
11	899.5	28.2	647	9	US-09-841-132-436
12	894	28.1	660	10	US-09-841-260-139
13	894	28.1	660	13	US-10-007-693-139
14	671.5	21.1	361	9	US-09-841-132-299
15	251	7.9	1463	9	US-09-971-536-69

16	235.5	7.4	1178	14	US-10-128-714-9240	Sequence 8240, Ap
17	234	7.3	1831	12	US-10-282-122A-71033	Sequence 71033, A
18	231.5	7.3	2478	9	US-09-815-242-5816	Sequence 5816, Ap
19	231.5	7.3	2478	9	US-09-815-242-12967	Sequence 12967, A
20	223.5	7.0	2481	12	US-10-282-122A-43762	Sequence 43762, A
21	218	6.8	603	12	US-10-282-122A-72004	Sequence 72004, A
22	217.5	6.8	6713	12	US-10-282-122A-43811	Sequence 43811, A
23	217	6.8	5795	9	US-09-815-242-12310	Sequence 12610, A
24	213	6.7	2434	9	US-09-815-242-5835	Sequence 5835, Ap
25	213	6.7	2434	9	US-09-815-242-12996	Sequence 12996, A
26	210	6.6	1122	12	US-10-282-122A-42783	Sequence 42783, A
27	210	6.6	1123	12	US-10-282-122A-70581	Sequence 70581, A
28	210	6.6	2271	12	US-10-282-122A-43924	Sequence 43924, A
29	208.5	6.5	4776	12	US-10-282-122A-73678	Sequence 73678, A
30	207.5	6.5	2344	9	US-09-815-242-12713	Sequence 12713, A
31	207	6.5	3533	12	US-10-282-122A-70177	Sequence 70177, A
32	204	6.4	2283	14	US-10-172-502-4	Sequence 4, Appl
33	203.5	6.4	2368	9	US-09-815-242-5635	Sequence 5635, Ap
34	203.5	6.4	2368	9	US-09-815-242-12389	Sequence 12389, A
35	201.5	6.3	1879	9	US-09-971-536-70	Sequence 70, Appl
36	201	6.3	459	12	US-10-282-122A-70453	Sequence 70453, A
37	201	6.3	1142	15	US-10-369-493-10778	Sequence 10778, A
38	199.5	6.3	2025	9	US-09-815-242-5703	Sequence 5703, Ap
39	199.5	6.3	3158	9	US-09-815-242-12611	Sequence 12611, A
40	199	6.2	6641	12	US-10-282-122A-70580	Sequence 70580, A
41	198.5	6.2	1215	9	US-09-815-242-5908	Sequence 5908, Ap
42	198.5	6.2	1289	9	US-09-815-242-13113	Sequence 13113, A
43	198	6.2	1222	12	US-10-282-122A-69799	Sequence 69799, A
44	196	6.1	560	9	US-09-815-242-13057	Sequence 13057, A
45	196	6.1	1924	12	US-10-282-122A-71846	Sequence 71846, A

ALIGNMENTS

RESULT 1
US-09-841-132-405
; Sequence 405, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:

; APPLICANT: Skeiky, Asir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 405
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae

Query Match 100.0%; Score 3187; DB 9; Length 651;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MWNPIGPGPIDETERTPPADLSAQGLEASANKSAQRIAGAEKPKESKTSVERWSI	60
Db	1	MWNPIGPGPIDETERTPPADLSAQGLEASANKSAQRIAGAEKPKESKTSVERWSI	60
Qy	61	LRSAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTDDYKTAQYDT	120
Db	61	LRSAVNALMSLADKLGIASSNSSSTSRSDVSTTATPTPPPTDDYKTAQYDT	120
Qy	121	IFTSTSLADIQAALVSLQDAVTNKTDAATDEETAAAEWETKNADAVKVAQITELAKY	180
Db	121	IFTSTSLADIQAALVSLQDAVTNKTDAATDEETAAAEWETKNADAVKVAQITELAKY	180
Qy	181	ASDNCAILDSLGLKTSFDLLQAALLQSVANNKAAELLKEMQDNFVPGKTPAIAQSLVD	240

Db 181 ASDNQAILDSGLKTSFDLLQALLOSQVANNKAAEILLKEMQDNVPVPGKTPALAQSLVD 240
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKALATAKTQIAEAQ 300
Db 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKALATAKTQIAEAQ 300
Qy 301 KKFPPSPILQEAQVQIAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Db 301 KKFPPSPILQEAQVQIAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Qy 361 DDAENETASILMSGFRQMIHNTENPDSSQAQOELAAQARAAGDDSAALADAOQ 420
Db 361 DDAENETASILMSGFRQMIHNTENPDSSQAQOELAAQARAAGDDSAALADAOQ 420
Qy 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Db 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Qy 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALAV 540
Db 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALAV 540
Qy 541 ISGNSRTLGDVYSQVSALQSVNQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Db 541 ISGNSRTLGDVYSQVSALQSVNQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Qy 601 STQKFTAKLESIFAEGRSRTAAEIKALSFEFNSLFIQOVLVNGISLYSGYLQ 651
Db 601 STQKFTAKLESIFAEGRSRTAAEIKALSFEFNSLFIQOVLVNGISLYSGYLQ 651

RESULT 2
US-09-564-479-2
; Sequence 2, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR FILING DATE: 1999-05-03
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-2

Query Match 100.0%; Score 3187; DB 10; Length 651;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDVERWSI 60
Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDVERWSI 60
Qy 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTATAPTPPPPTFDYKTOQATAYDT 120
Db 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTATAPTPPPPTFDYKTOQATAYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180

Qy 181 ASDNQAILDSGLKTSFDLLQALLOSQVANNKAAEILLKEMQDNVPVPGKTPALAQSLVD 240
Db 181 ASDNQAILDSGLKTSFDLLQALLOSQVANNKAAEILLKEMQDNVPVPGKTPALAQSLVD 240
Qy 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKALATAKTQIAEAQ 300
Db 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKALATAKTQIAEAQ 300
Qy 301 KKFPPSPILQEAQVQIAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Db 301 KKFPPSPILQEAQVQIAEKDLKNIKPADGSDVNPFGTTVGGSKQOQSSIGSRVSMML 360
Qy 361 DDAENETASILMSGFRQMIHNTENPDSSQAQOELAAQARAAGDDSAALADAOQ 420
Db 361 DDAENETASILMSGFRQMIHNTENPDSSQAQOELAAQARAAGDDSAALADAOQ 420
Qy 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Db 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKT 480
Qy 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALAV 540
Db 481 QISAGYDAYKSIDAYGRANDATRDVINNVSTPALTRSPRARTTEARGPEKTDQALAV 540
Qy 541 ISGNSRTLGDVYSQVSALQSVNQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Db 541 ISGNSRTLGDVYSQVSALQSVNQIIQSNPQANNEEIRQKLTSAVTKPPQFGYPYVQLSND 600
Qy 601 STQKFTAKLESIFAEGRSRTAAEIKALSFEFNSLFIQOVLVNGISLYSGYLQ 651
Db 601 STQKFTAKLESIFAEGRSRTAAEIKALSFEFNSLFIQOVLVNGISLYSGYLQ 651

RESULT 3
US-09-841-260-78
; Sequence 78, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Stromberg, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 78
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-78

Query Match 100.0%; Score 3187; DB 10; Length 651;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDVERWSI 60
Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASAANKSAEQAQRIAGAAKPKESKTDVERWSI 60
Qy 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTATAPTPPPPTFDYKTOQATAYDT 120
Db 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTATAPTPPPPTFDYKTOQATAYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
Qy 181 ASDNQAILDSGLKTSFDLLQALLOSQVANNKAAEILLKEMQDNVPVPGKTPALAQSLVD 240
Db 181 ASDNQAILDSGLKTSFDLLQALLOSQVANNKAAEILLKEMQDNVPVPGKTPALAQSLVD 240

QY 241 QTDAVATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db |||||
QY 241 QTDAVATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db |||||
QY 301 KKFPPSPILQEAQBMVIOAEKDLKNIKPADGSDVNPFGTTVGSKQOGSSIGSIRVSMLL 360
Db |||||
QY 301 KKFPPSPILQEAQBMVIOAEKDLKNIKPADGSDVNPFGTTVGSKQOGSSIGSIRVSMLL 360
Db |||||
QY 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSAALADAQK 420
Db |||||
QY 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSAALADAQK 420
Db |||||
QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASISGSSVKQLYKTSKSTGSDYKT 480
Db |||||
QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASISGSSVKQLYKTSKSTGSDYKT 480
Db |||||
QY 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTTRSPRARTFARGPEKTDQALARV 540
Db |||||
QY 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTTRSPRARTFARGPEKTDQALARV 540
Db |||||
QY 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
Db |||||
QY 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
Db |||||
QY 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
Db |||||
QY 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
Db |||||

RESULT 4
US-10-007-693-78
; Sequence 78, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 78
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-78

Query Match 100.0%; Score 3187; DB 13; Length 651;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNPPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQIAEAKPKESKTSVERWSI 60
Db |||||
QY 61 LRSVAVNALMSLADKLGIIASSNSSSTSRSDVDSTTATPTPPPTDDYKTKQAQTAIDT 120
Db |||||
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAADEETAAIAEWEETKNADAVKGAQITELAKY 180
Db |||||
QY 181 ASDNQATILDSIGKLTSPDLLQAALLOSVAANNKAAELLKEMQDNPVVGKTPALQAQSLVD 240
Db |||||
QY 241 QTDATATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db |||||

QY 301 KKFPPSPILQEAQBMVIOAEKDLKNIKPADGSDVNPFGTTVGSKQOGSSIGSIRVSMLL 360
Db |||||
QY 301 KKFPPSPILQEAQBMVIOAEKDLKNIKPADGSDVNPFGTTVGSKQOGSSIGSIRVSMLL 360
Db |||||
QY 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSAALADAQK 420
Db |||||
QY 361 DDAENETASILMSGFRQMIHMENTENPDSSQAQOELAAQARAARAAAGDDSAALADAQK 420
Db |||||
QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASISGSSVKQLYKTSKSTGSDYKT 480
Db |||||
QY 421 ALEAALGKAGQOQGIILNALGQIASAAVVSAGVPPAAASISGSSVKQLYKTSKSTGSDYKT 480
Db |||||
QY 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTTRSPRARTFARGPEKTDQALARV 540
Db |||||
QY 481 QISAGYDAYKSINDAYGRANDATROVINNVSTPALTTRSPRARTFARGPEKTDQALARV 540
Db |||||
QY 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
Db |||||
QY 541 ISGNSRTLGDVYSQVSALOSVMQIIQSNPOANNEEIRQKLTSAVTKPPQFGYPPVQLSND 600
Db |||||
QY 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
Db |||||
QY 601 STQKFIKLESLEFAGSRTAAEIKALSFTETNSLFIQOVLVNIIGSLYSGYLQ 651
Db |||||

RESULT 5
US-10-312-273-13
; Sequence 13, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P02503SWO
; CURRENT APPLICATION NUMBER: US/10/312,273
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019368.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 13
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-13

Query Match 100.0%; Score 3187; DB 15; Length 651;
Best Local Similarity 100.0%; Pred. No. 4e-225;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNPPIGPGIDETERTPPADLSAQGLEASAANKSAEQAQIAEAKPKESKTSVERWSI 60
Db |||||
QY 61 LRSVAVNALMSLADKLGIIASSNSSSTSRSDVDSTTATPTPPPTDDYKTKQAQTAIDT 120
Db |||||
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAADEETAAIAEWEETKNADAVKGAQITELAKY 180
Db |||||

Db 121 IFTSTSLADIQAALVSLQDAVTNINKDTAATAAETKADAVKVGQAQITELAKY 180
Qy 181 ASDNCAILDSLGKLTSPDLLQAALLOSANNKKAELLKEMQDPVVPKTPAQAQSLVD 240
Db 181 ASDNCAILDSLGKLTSPDLLQAALLOSANNKKAELLKEMQDPVVPKTPAQAQSLVD 240
Qy 241 QTDATATQIEKDGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
Db 241 QTDATATQIEKDGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQ 300
Qy 301 KKFDPSPILQBAEQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMML 360
Db 301 KKFDPSPILQBAEQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMML 360
Qy 361 DDAENETASILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAK 420
Db 361 DDAENETASILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAK 420
Qy 421 ALEAALGKAGQOQGIILNAGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
Db 421 ALEAALGKAGQOQGIILNAGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKT 480
Qy 481 QISAGYDAYKSNIDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARV 540
Db 481 QISAGYDAYKSNIDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARV 540
Qy 541 ISGNRTIGDVYSQVSAQSVMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
Db 541 ISGNRTIGDVYSQVSAQSVMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSND 600
Qy 601 STQKFIKLESIFAGSRTAAIBKALSPETNSLFTQQVLVNTIGSLYSGYLQ 651
Db 601 STQKFIKLESIFAGSRTAAIBKALSPETNSLFTQQVLVNTIGSLYSGYLQ 651

RESULT 6

US-09-564-479-4

; Sequence 4, Application US/09564479

; Publication No. US2003095973A1

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: WOMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; FILE REFERENCE: 032931/0230

; CURRENT APPLICATION NUMBER: US/09/564,479

; PRIOR FILING DATE: 2000-05-03

; PRIOR FILING DATE: 1999-05-03

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 4

; LENGTH: 583

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-564-479-4

Query Match 89.2%; Score 2844; DB 10; Length 583;

Best Local Similarity 99.8%; Pred. No. 5e-200;

Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 69 MSADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTOAQATYDTTFTSTSLA 128
Db 1 MSADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTOAQATYDTTFTSTSLA 60
Qy 129 DIQAALVSLQDAVTNINKDTAATAAETKADAVKVGQAQITELAKYSDNQAIL 188
Db 61 DIQAALVSLQDAVTNINKDTAATAAETKADAVKVGQAQITELAKYSDNQAIL 120

Qy 189 DSLGKLTSPDLLQAALLOSANNKKAELLKEMQDPVVPKTPAQAQSLVDQTDATQ 248
Db 121 DSLGKLTSPDLLQAALLOSANNKKAELLKEMQDPVVPKTPAQAQSLVDQTDATQ 180
Qy 249 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQKFFPDSPI 308
Db 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEBAQKFFPDSPI 240
Qy 309 LQBAEQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMMLDDAENETA 368
Db 241 LQBAEQVMIQAEKDLKNIKPADGSDVENPGTTVGGSKQOQSSISIRVSMMLDDAENETA 300
Qy 369 SILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAKALEAALGK 428
Db 301 SILMSGFRMHMENTENPDPSQAAQELAAQAAKAAAGDSDSAAALADAKALEAALGK 360
Qy 429 AGQOQGIILNAGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKTQISAGYDA 488
Db 361 AGQOQGIILNAGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKTQISAGYDA 420
Qy 489 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 548
Db 421 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 480
Qy 549 GDVYSQVSAQSVMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 608
Db 481 GDVYSQVSAQSVMQIITQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 540
Qy 609 LESLFAEGSRTAAIBKALSPETNSLFTQQVLVNTIGSLYSGYLQ 651
Db 541 LESLFAEGSRTAAIBKALSPETNSLFTQQVLVNTIGSLYSGYLQ 583

RESULT 7

US-09-564-479-8

; Sequence 8, Application US/09564479

; Publication No. US2003095973A1

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: WOMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; FILE REFERENCE: 032931/0230

; CURRENT APPLICATION NUMBER: US/09/564,479

; PRIOR FILING DATE: 2000-05-03

; PRIOR FILING DATE: 1999-05-03

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 8

; LENGTH: 490

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-564-479-8

Query Match 69.3%; Score 2208; DB 10; Length 490;

Best Local Similarity 99.8%; Pred. No. 1.8e-153;

Matches 453; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVNPIGPGPIDETETPPADLSAQGLEASAAKSAEAFQIACAEAKPKESKTDSEVERSI 60
Db 1 MVNPIGPGPIDETETPPADLSAQGLEASAAKSAEAFQIACAEAKPKESKTDSEVERSI 60
Qy 61 LRSVNALMSLADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTOAQATYDT 120
Db 61 LRSVNALMSLADKLGIASSSSSTSSADVSDSTATAPTPPPPTFDYKTOAQATYDT 120
Qy 121 IFTSTSLADIQAALVSLQDAVTNINKDTAATAAETKADAVKVGQAQITELAKY 180

Db 121 IFTSTSLADIQAALVSLQDAVTNKTAAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 Qy 181 ASDNQAILDSLGKLTSPDLQAALQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 Db 181 ASDNQAILDSLGKLTSPDLQAALQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 Qy 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Qy 301 KKPDPSPILQEAQWVIQAELKXNIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 Db 301 KKPDPSPILQEAQWVIQAELKXNIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAKAGDSDSAAALADAQK 420
 Db 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGVPP 454
 Db 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGVLP 454

RESULT 8

US-09-564-479-6

; Sequence 6, Application US/09564479

; Publication No. US20030095973A1

; GENERAL INFORMATION:

; APPLICANT: MURDIN, ANDREW D.

; APPLICANT: WOMEN, RAYMOND P.

; APPLICANT: WANG, JOE

; APPLICANT: DUNN, PAMELA

; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND

; TITLE OF INVENTION: USES THEREOF

; FILE REFERENCE: 032931/0230

; CURRENT APPLICATION NUMBER: US/09/564,479

; PRIOR APPLICATION NUMBER: 2000-05-03

; PRIOR FILING DATE: 1999-05-03

; PRIOR APPLICATION NUMBER: 60/141,276

; PRIOR FILING DATE: 1999-06-30

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 452

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-09-564-479-6

Query Match 69.2%; Score 2204; DB 10; Length 452;
 Best Local Similarity 100.0%; Pred. No. 3.le-153;
 Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEORAGAEAKPKESKTDSEVERWSI 60
 Db 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEORAGAEAKPKESKTDSEVERWSI 60
 Qy 61 LRSVAVNALMSLADKLGIASNSSSSSTSRSDVSTTATPTPPPTFDDYKTAQTAAYDT 120
 Db 61 LRSVAVNALMSLADKLGIASNSSSSSTSRSDVSTTATPTPPPTFDDYKTAQTAAYDT 120
 Qy 121 IFTSTSLADIQAALVSLQDAVTNKTAAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 Db 121 IFTSTSLADIQAALVSLQDAVTNKTAAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 Qy 181 ASDNQAILDSLGKLTSPDLQAALQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 Db 181 ASDNQAILDSLGKLTSPDLQAALQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 Qy 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Db 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300

Qy 301 KKPDPSPILQEAQWVIQAELKXNIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 Db 301 KKPDPSPILQEAQWVIQAELKXNIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAKAGDSDSAAALADAQK 420
 Db 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAKAGDSDSAAALADAQK 420
 Qy 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452
 Db 421 ALEAALGKAGQOQGGILNALGQIASAAVVSAGV 452

RESULT 9

US-10-289-762-776

; Sequence 776, Application US/10289762

; Publication No. US20040006218A1

; GENERAL INFORMATION:

; APPLICANT: Griffiths, R.

; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments

; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention

; TITLE OF INVENTION: and treatment of infection

; FILE REFERENCE: 9710-003-999

; CURRENT APPLICATION NUMBER: US/10/289,762

; CURRENT FILING DATE: 2003-03-27

; NUMBER OF SEQ ID NOS: 6849

; SEQ ID NO 776

; LENGTH: 478

; TYPE: PRT

; ORGANISM: Chlamydia pneumoniae

US-10-289-762-776

Query Match 68.5%; Score 2182; DB 15; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.4e-151;
 Matches 447; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEORAGAEAKPKESKTDSEVERWSI 60
 Db 3 MWNPIGPGPIDETERTPPADLSAQGLEASANKSAEORAGAEAKPKESKTDSEVERWSI 62
 Qy 61 LRSVAVNALMSLADKLGIASNSSSSSTSRSDVSTTATPTPPPTFDDYKTAQTAAYDT 120
 Db 63 LRSVAVNALMSLADKLGIASNSSSSSTSRSDVSTTATPTPPPTFDDYKTAQTAAYDT 122
 Qy 121 IFTSTSLADIQAALVSLQDAVTNKTAAATDEETAIAAEWETKNADAVKVGQAQITELAKY 180
 Db 123 IFTSTSLADIQAALVSLQDAVTNKTAAATDEETAIAAEWETKNADAVKVGQAQITELAKY 182
 Qy 181 ASDNQAILDSLGKLTSPDLQAALQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 240
 Db 183 ASDNQAILDSLGKLTSPDLQAALQSVANNKAAELLKEMQDNVPVPGKTPAIAQSLVD 242
 Qy 241 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
 Db 243 QTDATATQIEKDGNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 302
 Qy 301 KKPDPSPILQEAQWVIQAELKXNIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 360
 Db 303 KKPDPSPILQEAQWVIQAELKXNIPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMML 362
 Qy 361 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAKAGDSDSAAALADAQK 420
 Db 363 DDAENETASILMSGFRQMIHMFNTENPDSSQAQOELAAQAAKAGDSDSAAALADAQK 422
 Qy 421 ALEAALGKAGQOQGGILNALGQIASAAV 447
 Db 423 ALEAALGKAGQOQGGILNALGQIASAAV 449

RESULT 10

US-10-289-762-775

; Sequence 775, Application US/10289762

; Publication No. US20040006218A1

```

; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; thereof and uses thereof, in particular for the diagnosis, prevention
; and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 775
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-10-289-762-775

Query Match      31.4%; Score 1000; DB 15; Length 212;
Best Local Similarity 99.5%; Pred. No. 2.1e-65;
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 448 VSAGVPPAAASSIGSSVQLYKTSKSTGSDYKQTSAGVDAYKSIINDAYGRANDATRDV 507
Db 9 VSAGVPPAAASSIGSSVQLYKTSKSTGSDYKQTSAGVDAYKSIINDAYGRANDATRDV 68
QY 508 INNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTGLGVYQVSALQSVMOILOS 567
Db 69 INNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTGLGVYQVSALQSVMOILOS 128
QY 568 NPOANNEEROKLTSATVTKPPQFGYPVOLNSDSTQKFIKLESFAEGSRTAAAEIKALS 627
Db 129 NPOANNEEROKLTSATVTKPPQFGYPVOLNSDSTQKFIKLESFAEGSRTAAAEIKALS 188
QY 628 FETNSLFIQQLVNLGSLYSGYLQ 651
Db 189 FETNSLFIQQLVNLGSLYSGYLQ 212

RESULT 11
US-09-841-132-436
; Sequence 436, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Sreelky, Rasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 436
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
; US-09-841-132-436

Query Match      28.2%; Score 899.5; DB 9; Length 647;
Best Local Similarity 35.1%; Pred. No. 2.6e-57;
Matches 234; Conservative 118; Mismatches 262; Indels 53; Gaps 16;

QY 2 VNTPGPIDETETPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTDVSVERWSIL 61
Db 15 MNPIINGQI-----ASNSETKESKEA-----SFSASSVSSWSFL 52
QY 62 RSVNMLSLADKLGIIASSNSSSTSRs-ADVDSSTATATPTPPPTFDYKTAQTAAYDT 120
Db 53 SSAKHALISLRD--AILNKNSPPTDSLSQLEASTSTSTVTRVAARDYNEAKSNFDTAKSG 110
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEE--TATAAEWETKNADAVKVGAGITELAK 179
Db 111 LENAATTIAEYETKMAADLMAALQDMERLAKQAEVTRIKALQEKQ-----EVIDKLNQVVK 166
```

```

QY 180 YASDNQAILDSGLKLTSPDLLOALLOSVANNKKAELLKEMQDNPPVPGTKPTAQAOSLV 239
Db 167 LEKQNTKLTETLTFTTSDAQPAINSQLEIKNSADQIKOLEGQNI---SYEAVLTNAG 223
QY 240 DQTDATATQIEKDGNAIRDVAFAGNAGSAGVENAKSNNSISNIDSAKAAIATAKTQIAEA 299
Db 224 EVIKASSAGIKLGOALQSIIVDAGQSQAAVLOAQONNSPDNIATKLLIDAEIKVNEL 283
QY 300 QXK---FPDSPILOFAEQWVIAEKDLKNIKPADGSDVPNPOTTVGGSKQSSSISGI-- 354
Db 284 KQEHGLTDSPLVKAEEQISQAQKDIQBIKPS--GSDIPVGPS--GSAASAGSAGVALKS 341
QY 355 -----RVSMILLDDAENETASILMSFROMIHMFENTENPDSQAQQLAQAARAK--AA 406
Db 342 SNNSGRI SLLDDVDNEMAAIAMQGRSMIEOFVNNNPATAKELQAMEAQLTAMSDQLVG 401
QY 407 GDSAAALADAAQKALEAALGKAGQQQGLNALGQIASAAVVSAGVPPAAASSIGSSVKQ 466
Db 402 ADGEIAPBIAIKDALAQAL-KQPTDGLATAMGQVAFAAAKVGGSGAGTAGTGVQMNVKQ 460
QY 467 LYKT--SKSTGSDYKQTSAGVDAYKSIINDAYGRANDATRDVINNVSTPALTRSPVPRAR 524
Db 461 LYKTAFTSTSSSYAAALSDDGYSAYKTLNLSYSESRs-GVQSAISQTANPALSRVSRSRG 519
QY 525 TEARG-PEKTDQALARVISGNSRTGLGVYQVSALQSVMOILOSPOANNEEROKLTS 583
Db 520 IESQGRSADASQRAAETIVRDSQTGLGVYSLRLQVLDLSLMTSTIVSNPQVNOEIMQKLTR 579
QY 584 VTKPPQFGYPVOLNSDSTQKFIKLESFAEGSRTAAAEIKALSFPETNSLFIQQLVNLG 643
Db 580 ISKAPQFGYPAVONSADSLQKFAAQLEREFVDGERSLAESRENAFRKQPAFIQQLVNLTA 639
QY 644 SLVSGYL 650
Db 640 SLFSGYL 646

RESULT 12
US-09-841-260-139
; Sequence 139, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 139
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
; US-09-841-260-139

Query Match      28.1%; Score 894; DB 10; Length 660;
Best Local Similarity 34.6%; Pred. No. 6.8e-57;
Matches 233; Conservative 126; Mismatches 262; Indels 52; Gaps 17;

QY 6 GPGPIDETETPPADLSAQGLEASANKSAEQAQRIAGAEAKPKESKTDVSVERWSILRS 65
Db 11 GPESVSSNOSNMPIINGQIASNSETKESKEA-----SEASP--SASSSVSSWSFLSSAK 63
QY 66 NALMSLADKLGIIASSNSSSTSRs-ADVDSSTATATPTPPPTFDYKTAQTAAYDTFTS 124
Db 64 NALISLRD--AILNKNSPPTDSLSQLEASTSTSTVTRVAARDYNEAKSNFDTAKSLENA 121
QY 125 TSLADTQAALVSLQDAVTNIKDTAATD-----EET-----ATAAEWETKNADAVKVG 173
Db 122 KTLAEYETKMAADLMAALQDMERLANSDPNNHTEVNNYNNKKALEAQKDTID-----K 173
```

[illegible]

```

RESULT 13
US-10-007-693-139
; Sequence 139, Application US10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 139
; LENGTH: 660
; TYPE: PRN
; ORGANISM: Chlamydia trachomatis
US-10-007-693-139

      Query Match      28.1%; Score 894; DB 13; Length 660;
      Best Local Similarity 34.6%; Pred. No. 6,8e-57;
      Matches 233; Conservative 126; Mismatches 262; Indels 52; Gaps 17

Qy      6  GPGPIDETERTPPADLSAQGLEAAGAAANKSAQAORIAGAEAKPKGSKTOSVERWSTLRSAV 65
Db      11  GPESVSSNQSNMNPINGQIASNSETKESTRA-----SEAGP--SASSSVSSWSFLSSAK 63

Qy      66  NALMSLADKLGIASSNSSSSSTSRs-ADVDSITATAPTPTPTFDYKTAQATYDTITS 124
Db      64  NALISLRD--AILKNSSPTSLSQLEASTSTSTVTRVAAKDYDEAKSNFTAKSGLENA 121

Qy      125  TSLADIOAALVSLQDAVNIKDTATD-----EET-----AIAAEWETKNADAVKVGQAQ 173
Db      122  KTLAEYETKMADLMAALQOMERLANSDPSNNHTEVNNIKKALEAQKOTID-----K 173

Qy      174  ITELAKYASDNCAILDSLGKLTISFDLQAAILQSVAANNKKAELLKEMQDNVPVPGKTPA 233

```

Db	174	LNKLVLQNKQSLTEVLKTTDSADQIPALNSQLEINKNSSADQILKOLERNQI---SYEA	233
Qy	234	IAQSLVNDQTDATATQIEKGNAINRDAYFAGNAGSAGVENAKSNNSIENIDSAKAAIATAK	293
Db	231	VLTNAGEVIRKASSEAGIKLGCALQSIVDAGDQSQAAVLQAQNNSPDNIAATKELIDAAE	290
Qy	294	TQTABAQK---PPDSPIQEAQOMVLOAEKDLKNIKPADGSDVNPCTTTVGGSKQGGSS	350
Db	291	TKYNELKQHEHTGLTSDPLVKKABEQISQAQKIQIEIKPS-GSDPIIVGFS-GSAASAGSA	348
Qy	351	IGSI-----RVSMLLDDAENETASILMSGFRQIMHMFNTENPDQSQAQOQELAAQARAA	403
Db	349	AGALKSSNNSGRISLLDDVDNENKAATALQGFRSKIEQFNVNNPATAKELQAMEAQLTAM	408
Qy	404	K---AAGDDSAALADAQKALEAALGKAGQOQOGLNALGQIASAAVVSVAGVPPAAASSI	460
Db	409	SDQLVGADGELPABEQIAIKDALAQAL-KQPSADGLATAMGQVFAAAKVGGSAGTAGTV	467
Qy	461	GSSVKQLYKT-SKSTGSDYKTOISAGYDAYKSNIDAYGRAENDATRDVINNVSPALTR	518
Db	468	QMNVKQLYKTAFTSTSSSSAAAALSDGYSAYKTLNLSYSEBS-GVQSAISQGTANFALSR	526
Qy	519	SVPRARTEARG-PEKTPQALARVIGSGNSRTLGDVYSQVSAALQSVNQIITQSPQANNEIR	577
Db	527	SVSRSGTESQGRSADASQRAAETIVRDSQTLGDVYSRLQVLDSLMSTIVSNPQANQBEIM	586
Qy	578	QKITSVATKPPQFGYPYVQIJSNDSTQKPIAKLESFABGSGTAARIKALSTETNSLFTQQ	637
Db	587	QKLTASISKAPQFGYPVQNSADSLQKFAAQULEREFEVDGERSLAEBSQENAFRKQPAFTQQ	646
Qy	638	VLVNIIGLSGYL 650	
Db	647	VLVNIASLFSGYL 659	
RESULT 14			
US-09-841-132-299			
; Sequence 299, Application US/09841132			
; Patent No. US2002006184A1			
; GENERAL INFORMATION:			
; APPLICANT: Bhatia, Ajay			
; APPLICANT: Skeiky, Yasir A.W.			
; APPLICANT: Probst, Peter			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND			
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION			
; FILE REFERENCE: 210121.469C8			
; CURRENT APPLICATION NUMBER: US/09/841.132			
; CURRENT FILING DATE: 2001-04-23			
; NUMBER OF SEQ ID NOS: 599			
; SOFTWARE: FastSeq for Windows Version 3.0/4.0			
; SEQ ID NO 299			
; LENGTH: 361			
; TYPE: PRT			
; ORGANISM: Chlamydia			
US-09-841-132-299			
Query Match 21.1%; Score 671.5; DB 9; Length 361;			
Best Local Similarity 42.4%; Pred. No. 5.9e-41;			
Matches 154; Conservative 62; Mismatches 130; Indels 17; Gaps 8			
Qy	301	KKFPDPSPIQEAQOMVLOAEKDLKNIKPADGSDVNPCTTTVGGSKQGGSSIGSI-----	354
Db	2	QEIADSPFLVKRAEQINQAQDIIQTIPS-GLDPIIVGFS-GSAASAGSAGALKSSNNS	59
Qy	355	-RVSMLLDDAENETASILMSGFRQIMHMFNTENPDQSQAQOQELAAQAPAAK---AAGDDS	410
Db	60	GRISILLDDVDNENKAATAMQGFRSKIEQFNVNNPATAKELQAMEAQLTAMSDQLVGADGE	119
Qy	411	AAALADAQKALEAALGKAGQOQOGLNALGQIASAAVVSVAGVPPAAASSIGSSVKQLYKT	470
Db	120	LPABEQIAIKDALAQAL-KQPSADGLATAMGQVFAAAKVGGSAGTAGTVQMNVKQLYKT	178

```

RESULT 14
US-09-841-132-299
; Sequence 299, Application US/09841132
; Patent No. US2002061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-299

Query Match          21.1%; Score 671.5; DB 9; Length 361;
Best Local Similarity 42.4%; Pred. No. 5.9e-41;
Matches 154; Conservative 62; Mismatches 130; Indels 17; Gaps 8

Qy      301  KKFPDSPLOAEQWVICAERKDLNKKPKADGSDVNPNGTITVGGSKQGSSIGSI----- 354
Db      2   QETADSPFLVKKAEQIQNAQDIIITPS-GLDIPITVGPS-GSAASAGSAGALKSSNNS 59

Qy      355  -RVSMLLDDRAENETASTILMSGFRQMIMHFNENTENPDSQAQQLAEAAQFAAK---AAGDDS 410
Db      60  GRISLLDDVDVNEAATAMQGFSEMIQEFNVNPNPATAKELQAMEAQLTAMSDQLVGADGE 119

Qy      411  AAAALADAQKALEAALGKAGQOQGLILNALGQIAGSAAVVSAGVPPAAASSIGSSVKQLYKT 470
Db      120  LPARIQAIKALAAQAL-KQPSADGLIATAMGQVFAAAKVGSGSAGTAGTGVQNNVKQLYKT 178

```

```
QY 471 --SKSTGSDYKTIQISAGYDAYKSIINDAYGRANDTRDVIINNVSTPALTRSVPRARTEAR 528
Db 179 AFSSTSSSSYAALSDGYKATLNSLYSESRG-GVQSAISQOTANPALSRSVSRGIESQ 237
QY 529 G-PEKTDQALARVISGNSSTLGDVYQVSOVALQSVQIIOSNPOANNEETROKLTSAVTKP 587
Db 238 GRSADASQAAETIVRDSQTLGDVYSRQLQVLDSLMSTIVSNPQANQOEIMQKLTASISKA 297
QY 588 PFGPYVYVQLSNDSTOKFTAKLESIPAEGRSTAAAEIKALSFETNSLFIOQVLVNTGSIYS 647
Db 298 PFGPYPAVONSVDLSQKFAQLEREFVQERSLAESQENAFKQKPAFIOQVLVNTIASLES 357
QY 648 GYL 650
Db 358 GYL 360

RESULT 15
US-09-971-536-69
; Sequence 69, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Bloksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul
; APPLICANT: Reid, Julian
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and Methods
; FILE OF INVENTION: Using Them
; FILE REFERENCE: 1043C2
; CURRENT APPLICATION NUMBER: US/09/971,536
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 1463
; TYPE: PRF
; ORGANISM: Lactobacillus rhamnosus
US-09-971-536-69

Query Match 7.9%; Score 251; DB 9; Length 1463;
Best Local Similarity 22.1%; Pred. No. 3.3e-09;
Matches 159; Conservative 106; Mismatches 268; Indels 186; Gaps 27;

QY 19 ADLSAQGLEAS-----AANKSAPQRIAGAAKPKESKT--DSVERWSIL-----RSAY 65
Db 720 ADVTASQAKATSAASDAASAYASEAQSIAGSHADNMEIKSLASDAEKQSIQIALAASKSAA 779
QY 66 NALMSLADKLGIASSNSSSSRSRSDVDTTATPTPPPTFFDDYKTOAQTA----- 117
Db 780 ASSSAAASAIVASASASASASAAAANVSNDASA-NSAAAAYDSYASEASASANDSSG 838
QY 118 YDT--IFTSTSLADIQALVSLQDAV-TNIKDTAATDEETAI--AAEWETKNADAVKGA 172
Db 839 YATASFAASAAAAMSAALSTAQVAAKAVVSDAAAAGSAAAVASAAQSDSKNKQATAATA 898
QY 173 -----QITELAKVADNQALDLSGLKLTSTFDLQALLQSVAN-----NNKA 214
Db 899 RSQALDDDLINKIKSLTVDYASGASSASEAGQASTATSAYASASSASASEAGSYAHQAGSSA 958
QY 215 AELLKEMQDNFVPGKTPAIA-----QSLVDQTDATATQIEKDGNAIRDAYFAG- 263
```

```
Db 959 SDAVQVQ-SGSAAOHASTAASAAASVYPKDSGTQISLASQAASEAAKASNAASATSAAYGF 1017
QY 264 QNASGAVENAKSNNSISNIDSAXAAIATAKTQIAEPAQKKFPDPSPILOBAEQMVIOAEKDL 323
Db 1018 SAASDASQAKTAASADVAVASSAASASTANSNAAAASATKAGDS----- 1060
QY 324 KNIKPADG--SDVPNPGTIVGSKQOQSSIGSIRVSMLLDDAENETASILMSGFRQMIHM 381
Db 1061 ---KAAAGFSSAAASAAASAKGAEAIVASEAASAAAS---DDSVASSAASAAAGFDK---- 1110
QY 382 FNTENPDQAAQOQELAAQARAARAAAGDDSAALADAAQKALEAALGKAGQOQGIILNALQ 441
Db 1111 -----AASRAEGRAASRAASRAASA-----AAGTRGSSSSASEAGQ 1148
QY 442 IASAAVVSAGVPPAAASSIGSVKQL-YKTSKSTG-----SDY 478
Db 1149 ASTATSVYASAAASSASEAGSYAHQAGSSASEATGHASSATSQAASAAASRYPSDSGI 1208
QY 479 KTOISAGYDAYKSIINDAYGRANDT-----TRDVINNVSTPALTRSV 520
Db 1209 QSDVSIASSAASTASGAASAAQSEASTASSAAASHASEQASIASSEDVSSSSAASVASSAA 1268
QY 521 PRARTEARGPEKTDQALARVISGNSRTLGDVYSQ-VSALQSYMQIIQS-----NPQANN 573
Db 1269 SAASSNAK-----AGNSSAAG-IYSHAASAAASAAASAKSAESQASAAASAAASD 1314
QY 574 BEIROKLTSAVTKPPQPGYPYVOLNDSTOKFTAKLESIPAEGRSTAAAEIKALSFETNS 632
Db 1315 DSVASSAASAA-----LSDDAKASSAADV-----ASSATTAATISSATSLADQS 1357
```

Search completed: March 24, 2004, 06:16:57
Job time : 153.132 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:39:25 ; Search time 11.9698 seconds
(without alignments)
2831.945 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187
Sequence: 1 MVNPIGPGIDERTPPAD.....SLFIQQLVNICSLYGYLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	210	6.6	1120	1 STFR_ECOLI	P76072 escherichia
2	187.5	5.9	2541	1 TLN1_HUMAN	Q9Y490 homo sapien
3	186.5	5.9	1306	1 MSB2_YEAST	P32334 saccharomyc
4	186	5.8	8545	1 ANCL1_CAEEL	Q9N4M4 caenorhabdi
5	182.5	5.7	1238	1 SBCC_RHOCA	O68032 rhodobacter
6	182.5	5.7	2541	1 TLN1_MOUSE	P26039 mus musculu
7	178	5.6	641	1 HRPK_PSESY	P41501 pseudomonas
8	177	5.6	2090	1 N214_HUMAN	P35858 homo sapien
9	176.5	5.5	545	1 HTR5_HALN1	Q48318 halobacteri
10	174.5	5.5	1140	1 YM96_YEAST	Q04893 saccharomyc
11	174	5.5	774	1 STF_LAMBD	P03764 bacterioph
12	172.5	5.4	3644	1 MINT_MOUSE	O62504 mus musculu
13	169	5.3	1528	1 SPAA_STRDO	P21979 streptococ
14	168.5	5.3	881	1 PRV2_YEAST	P47033 saccharomyc
15	168	5.3	810	1 HTR4_HALSA	Q48317 halobacteri
16	167.5	5.3	2492	1 TALA_DICDI	P54633 dictyosteli
17	167	5.2	810	1 HTR4_HALN1	Q9hp84 halobacteri
18	165.5	5.2	1562	1 SPAP_STRMU	P23504 streptococ
19	165.5	5.2	1565	1 PAC_STRMU	P11557 streptococ
20	163.5	5.1	2542	1 TLN2_HUMAN	Q9Y496 homo sapien
21	163	5.1	705	1 CWBA_BACSU	Q02113 bacillus su
22	160.5	5.0	1690	1 C190_DROME	Q9V7e5 drosophila
23	160	5.0	535	1 HTR1_HALN1	P33741 halobacteri
24	159.5	5.0	1411	1 TCOF_HUMAN	Q13428 homo sapien
25	158.5	5.0	1379	1 YFF9_SCHPO	Q14066 schizosacch
26	158	5.0	758	1 IMMT_HUMAN	Q16891 homo sapien
27	158	5.0	763	1 HTR2_HALN1	Q9hp81 halobacteri
28	158	5.0	764	1 HTR2_HALSA	P71410 halobacteri
29	157.5	4.9	1011	1 SCA4_RICAF	Q9aj83 rickettsia
30	157.5	4.9	1093	1 TMF1_HUMAN	P82094 homo sapien
31	157.5	4.9	1509	1 MYSN_ACACA	P05659 acanthamoeb
32	157.5	4.9	1849	1 IGAA_HAEIN	P45386 haemophilus
33	157.5	4.9	3164	1 TEGU_HSV11	P10220 herpes simp

RESULT 1

STFR_ECOLI

ID STFR_ECOLI STANDARD; PRT; 1120 AA.

AC P76072; P77560;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Side tail fiber protein homolog from lambdoid prophage Rac.

GN STFR OR B1372.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97251357; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

Ra Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

Ra Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12;

RX MEDLINE=97251357; PubMed=9097039;

RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,

Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,

Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,

Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,

Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,

Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;

RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome

corresponding to the 28.0-40.1 min region on the linkage map.";

RL DNA Res. 3:363-377(1996).

CC -! SIMILARITY: Belongs to the tail fiber family.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

34 157.5 4.9 5171 1 BPEA_HUMAN O94833 homo sapien
35 157 4.9 962 1 IF2_NEIMB Q9JYD2 neisseria m
36 156.5 4.9 1337 1 DEXT_STRDO P39653 streptococ
37 156 4.9 2843 1 APC_HUMAN P25054 homo sapien
38 155.5 4.9 1022 1 SCA4_RICCN Q52658 rickettsia
39 155 4.9 1161 1 DAN4_YEAST P47179 saccharomyc
40 155 4.9 1957 1 SPOF_SCHPO Q10411 schizosacch
41 154.5 4.8 757 1 IMMT_MOUSE Q8Cag8 mus musculu
42 154.5 4.8 962 1 IF2_NEIMA Q9Jtb5 neisseria m
43 154 4.8 626 1 CBIB_ECOLI P04479 escherichia
44 153.5 4.8 991 1 SCA4_RICSI O9aj77 rickettsia
45 153.5 4.8 2779 1 LVA_DROME Q8ms1 drosophila

ALIGNMENTS

```
KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
SQ SEQUENCE 1120 AA; 113779 MW; 542E59D71EE795B4 CRC64;

Query Match      6.6%; Score 210; DB 1; Length 1120;
Best Local Similarity 20.4%; Pred. No. 0.005;
Matches 154; Conservative 102; Mismatches 294; Indels 206; Gaps 25;

QY 8 GPIDETETPPA---DLSAQGLASAAKAEQRIAGAEAKPKESKTDVSVERWSILRS 63
DB 91 GAMEDDDARPALRRFELM---VEEVARNASAVQNTAAAKSKASDASTSAREATTAAD 147
QY 64 AVNALMSLADKLGIIASSNSSSTSRSDVDTTATPTPPPTDDYKTKQATYDITFT 123
DB 148 AADSARAASSTAGQAAS-SAQSSASSAGTATKATEASKSAAAAEESKSAATSGAAGT 206
QY 124 S--TSLADIQAALVSLQDAVNINIKDTAATDEETALAEWETKNADAVKVGQITELAKYA 181
DB 207 SETNASASLQSAATSASTATTKASE-AATSARDAAAKSEKAAKSE-----TNASSSA 257
QY 182 SDNQALIDSLGKLTFDLLQALLOSVAANNKKAELLKEMQDNVPVPGKTPALIAQSLVDQ 241
DB 258 SSAASSATAAGN-----SAKAATSETNARSSETAGQASAAAGSKTAAASASNAS 310
QY 242 TDAATQTEKGNAIRDAYFQGNAGAVENAKNSNISNTSDSKAAIATATKQIABAQK 301
DB 311 TSA-----GQASASATAAGKSA-----ESAASSASTATTKAGEATE 346
QY 302 KFPDPSILQAEQVQIQEKLNKIPADGSDVPNPGTTVGSKQGGSSIGSIRVSMLLD 361
DB 347 Q-----ASAARASASAAKTSETNAKASE-----TSAESSKTAAASSASNAS 388
QY 362 DAENETATILMSGFRQMIMENTNPDSQAQQLAAQARAQAAGDSDSAAAAALADAQKA 421
DB 389 -----SAS-----SASAKDEATROASAASAKSSATTASTAKATEAAGSA 425
QY 422 LEALGKAGQOQGIINALGQTAASAVSAGVPPRAASSIGSVKLYKTSKTSQSDYKQ 481
DB 426 TAAQSKSTAESAAATRAETAAKRAEDIASAVALDASTTKKGIVOL---SSATNSTSETL 482
QY 482 ISAGYDAVKSINDAYGRANDATRD-----VINNVSTPALT-----RSVPART 525
DB 483 AATP-----KAVKSAVDNAEKILQKQNGADIPDKGCFLLNNINAVSKTDFADKGRMYRV 538
QY 526 EA-----RGPEKTDQALARY-IGSNRTSLGD----- 550
DB 539 NAPAGATSGKYFVVVMSAGSVSELASRVITTTATRTAGDPMNCFNGFVMPGGWTR 598
QY 551 ---VYSQVSALQSVNQIICSNPQAN-NBEIRKQLTSVTKPKPQFGYPVQLS----- 598
DB 599 GRAYAGMFWQYQNNERAIHSINMWSNKGDDLSRVFYVDGAAPVFVAFIEDGLSISAPGADL 658
QY 599 -NDSTQKPIA-----KLESLEPAEG-----SRT 619
DB 659 VVNDITYKFGATNPATECIAADVILDFKSGRGFYBESHSLIYVNDNLSCKLFATDIEIVARG 718
QY 620 AAEIKALSFETNSLFIQO-----VLVNGISLYSGY 649
DB 7-9 GQOIRMIGGEYGALEWENDGAKTYLLLTQGGDYVGGW 754

RESULT 2
TLN1 HUMAN STANDARD; PRT; 2541 AA.
AC Q9Y490; Q9NZQ2; Q9UHH8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Talin 1.
GN TLN1 OR TLN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC
NCBI_TaxID=9606;

[1] SEQUENCE FROM N.A.
RP Mao L., Fan Y.H.;
RA "Complete cDNA sequence of human talin.";
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2] SEQUENCE FROM N.A.
RP MEDLINE=20079169; PubMed=10610730;
RA Ben-Yosef T., Francomano C.A.;
RT "Characterization of the human talin (TLN) gene: genomic structure, chromosomal localization, and expression pattern.";
RL Genomics 62:316-319(1999).
CC -!- FUNCTION: Probably involved in connections of major cytoskeletal structures to the plasma membrane. High molecular weight cytoskeletal protein concentrated at regions of cell-substratum contact and, in lymphocytes, at cell-cell contacts (By similarity).
CC -!- SUBUNIT: Binds with high affinity to vinculin and with low affinity to integrins (By similarity).
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -!- SIMILARITY: Contains 1 I/LWEQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
EMBL; AF078628; AAD13152.1; .
DR EMBL; AF177198; AAF23322.1; .
DR EMBL; AF178534; AAF27330.1; .
DR EMBL; AF178081; AAF27330.1; JOINED.
DR Genew; HGNC:11845; TLN1.
DR MIM; 186745; .
DR GO; GO:0005925; C:focal adhesion; NAS.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; NAS.
DR GO; GO:0006928; P:cell motility; NAS.
DR GO; GO:0007016; P:cytoskeletal anchoring; NAS.
DR InterPro; IPR000299; Band 4.1.
DR InterPro; IPR002558; ILWEQ.
DR Pfam; PF00373; Band 41; 1.
DR ProDom; PD011820; ILWEQ; 1.
DR SMART; SM00295; B41; 1.
DR SMART; SM00307; ILWEQ; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS00657; FERM_3; 1.
DR PROSITE; PS00945; I_LWEQ; 1.
DR PROSITE; PS00945; I_LWEQ; 1.
DR Structural protein; Cytoskeleton.
KW DOMAIN 86 403 FERM.
FT DOMAIN 2340 2533 I/LWEQ.
FT CONFLICT 824 824 S -> R (IN REF. 2; AAF27330).
FT CONFLICT 1227 1227 S -> L (IN REF. 1).
FT CONFLICT 1549 1549 P -> A (IN REF. 2; AAF27330).
FT CONFLICT 1604 1604 Q -> K (IN REF. 2; AAF27330).
FT CONFLICT 1701 1701 E -> Q (IN REF. 2; AAF27330).
FT CONFLICT 1718 1718 H -> N (IN REF. 2; AAF27330).
FT CONFLICT 1966 1966 A -> R (IN REF. 1).
FT CONFLICT 2256 2256 MISSING (IN REF. 2; AAF27330).
SQ SEQUENCE 2541 AA; 269717 MW; 47FD9B22BCF47296 CRC64;

Query Match      5.9%; Score 187.5; DB 1; Length 2541;
Best Local Similarity 20.0%; Pred. No. 0.17;
Matches 157; Conservative 121; Mismatches 319; Indels 187; Gaps 32;

QY 1 MVNPIGPGPIDETERTPPADLSAQGLE-----ASAANKSAEAPQRIAGAEAKPKESKTDVS 55
DB 722 VVATPISFVCQEQILVEAGRLVAKAVEGCVSAQAATEDGQLLRGVGAAA----- 771
QY 56 ERWSILRSVAVNALMSLADKLGIIASSNSSSTSRSDVDTTATPTPPPTDDYKTKQ 115
```



```
FT DOMAIN 5203 5274 COILED COIL (POTENTIAL).
FT DOMAIN 5339 5409 COILED COIL (POTENTIAL).
FT DOMAIN 5444 5560 COILED COIL (POTENTIAL).
FT DOMAIN 5638 5696 COILED COIL (POTENTIAL).
FT DOMAIN 5759 5879 COILED COIL (POTENTIAL).
FT DOMAIN 5938 6055 COILED COIL (POTENTIAL).
FT DOMAIN 6106 6177 COILED COIL (POTENTIAL).
FT DOMAIN 6242 6312 COILED COIL (POTENTIAL).
FT DOMAIN 6347 6463 COILED COIL (POTENTIAL).
FT DOMAIN 6541 6599 COILED COIL (POTENTIAL).
FT DOMAIN 6662 6782 COILED COIL (POTENTIAL).
FT DOMAIN 6841 6958 COILED COIL (POTENTIAL).
FT DOMAIN 7009 7080 COILED COIL (POTENTIAL).
FT DOMAIN 7145 7215 COILED COIL (POTENTIAL).
FT DOMAIN 7250 7366 COILED COIL (POTENTIAL).
FT DOMAIN 7444 7502 COILED COIL (POTENTIAL).
FT DOMAIN 7565 7685 COILED COIL (POTENTIAL).
FT DOMAIN 7744 7861 COILED COIL (POTENTIAL).
FT DOMAIN 7912 7983 COILED COIL (POTENTIAL).
FT DOMAIN 8048 8118 COILED COIL (POTENTIAL).
FT DOMAIN 8153 8204 COILED COIL (POTENTIAL).
FT DOMAIN 8273 8329 COILED COIL (POTENTIAL).
FT DOMAIN 8370 8390 COILED COIL (POTENTIAL).
FT DOMAIN 8486 8545 KLARSICHT.
FT DOMAIN 1655 1658 POLY-ARG.
FT DOMAIN 3021 3029 POLY-LYS.
FT DOMAIN 3924 3932 POLY-LYS.
FT DOMAIN 4878 4886 POLY-LYS.
FT DOMAIN 5781 5789 POLY-LYS.
FT DOMAIN 5684 5692 POLY-LYS.
FT DOMAIN 7587 7595 POLY-LYS.
SQ SEQUENCE 8545 AA; 9566470 MW; 02A94D994BEE19E3 CRC64;

Query Match
Best Local Similarity 20.9%; Pred. No. 0.94; Length 8545;
Matches 155; Conservative 121; Mismatches 253; Indels 214; Gaps 34;

Qy 2 VNPFGPDIETERTPPADLSAQGLEASAA-----NKSAEQAQRIAGAAKPKESKTDVER 57
Db 4456 VSGVKPDELDDKERAENYNDLLAR--LATAADVLNKAELQVAKADEK-SLHDSVDR 4512
Qy 58 WSIILRSVAVNMLSLADKLGIASSNSSSTSFSAVDSTATAPTEPPTPDYKTAQTA 117
Db 4513 -----IVSRVLPL-----VRESDELHNNAEAVPTQYAPKABELKEVEAA 4552
Qy 118 YDTIFTSTS-----LADIAQAALVSLQDAVNTKDTA-----ATDEETAJAAEWETKNAD 166
Db 4553 KAVIANAPSSDAHVQQLQEQAVATAETLIPDLERARLWNEFLAARNIDIALIEQLQQPLD 4612
Qy 167 AVKGAQITELAKYAS-DNQAILDSLGKLTSLFDLLQALLOSVAANNKKAELLKEMQNP 225
Db 4613 AVL--AQPKRSABEAQAQVENLRNNSQQLSLD-----NKIANLQRISELDPLES-- 4661
Qy 226 VVPCKTPALQSLVDQDTATATQTEKD-----GNAIRDVAFAGQNASGA----- 269
Db 4662 -----AYADVRFEDVDVAEOTRQYDDVLDNDVAARELEDETLKQSASQVANEIDDISKM 4714
Qy 270 VENAKSNNSI-----SNIDSAKAAIATAKTI--AEAQKFP--PDGFIQAEQWVIQA 319
Db 4715 IDSTDPSERSILDTIAKSDIPALKAQINRIKORIVNADASRKHVTTDPKIAEDLDNKLAKL 4774
Qy 320 BKDLKN-IKPDGSDVNPNGTVGSKQGSISIRVSMLLDADENETASILMSGRFOM 378
Db 4775 QTELDDAITSDEHD-----KEQ-----LILSLKLN-----ISQFEQ- 4806
Qy 379 IHMFNTENPDQAQOELA-----AQABAQAAGDSDSAABA----- 414
Db 4807 IPLDOLKSDDLTKTAEKTEITNSLKPEAPEPLAKIQELREAKRVGDEARSAHDQIVALEX 4866
Qy 415 LADAQKALEAALGKAGQOQGIINALGQIASAAVVSAG--VPPAAASSTGSSV----- 464
Db 4867 EADVTAKESAKKKKKKSPQEMIDELSAKVEAKALIPKIEBAKKNENLPADDKKA 4926
```

```
Qy 465 KOLYKTSKSTGSDYKTKQISAGVDAYKSIDA-----YGRANDATRDVNNVSTPALT- 517
Db 4927 EQLVSNLEAFVKDVEQVSEKQDELKLNNDANDAIKRLGDALDDAKTKVPS-SVPALSE 4985
Qy 518 ---RSVPR-----ARTEARGPEKTDQALARVSGNSRTL----- 548
Db 4986 FKDIRIAPHLATLVEAVNDVPASPVSVALRDRA-AKFTSDLEKNTQKTGDDDEKRADELK 5044
Qy 549 GDVYSQVSALQSYMQLIQSNPQANNEEIRKLTSAVTKPPQFGYPYVQLSNSTQKFIK 608
Db 5045 NDVGNVAKNVEDVSKYQNPQP-----LDVAKDDANKLKAT 5081
Qy 609 LESL--FAEGS-----RTAAEIK 624
Db 5082 VEQLTKLAESSDKIDPQVAKDIK 5104

RESULT 5
SBCC_RHOCA STANDARD; PRT; 1238 AA.
ID SBCC_RHOCA
AC O68032;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nuclease sbCCD subunit C.
GN SBCC.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB1003 / St Louis;
RA MEDLINE=97404404; PubMed=9256491;
RX Vileek C., Paces V., Maltsev N., Paces J., Haselkorn R., Fonstein M.;
RT "Sequence of a 189-kb segment of the chromosome of Rhodobacter
capsulatus SB1003 ";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9384-9388(1997).
CC -!- FUNCTION: SbcCD cleaves DNA hairpin structures. These structures
can inhibit DNA replication and are intermediates in certain DNA
recombination reactions. The complex acts as a 3'->5' double
strand exonuclease that can open hairpins. It also has a 5'
single-strand endonuclease activity (By similarity).
CC -!- SUBUNIT: Heterodimer of sbCC and sbCD (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. SbcC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by, and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: AF010496; AAC16118.1; -.
DR PIR: T03465; T03465.
DR InterPro: IPR003439; ABC_transporter.
KW Hydrolase; Nuclease; Exonuclease; Endonuclease; DNA replication;
KW DNA recombination; ATP-binding; Coiled coil.
FT NP BIND 37 44 ATP (POTENTIAL).
FT DOMAIN 395 438 COILED COIL (POTENTIAL).
FT DOMAIN 466 487 COILED COIL (POTENTIAL).
FT DOMAIN 521 600 COILED COIL (POTENTIAL).
FT DOMAIN 724 770 COILED COIL (POTENTIAL).
FT DOMAIN 901 943 COILED COIL (POTENTIAL).
FT DOMAIN 1019 1052 COILED COIL (POTENTIAL).
SQ SEQUENCE 1238 AA; 128046 MW; 2B6BC6C63A859AE1 CRC64;

Query Match
Best Local Similarity 5.7%; Score 182.5; DB 1; Length 1238;
Matches 155; Conservative 116; Mismatches 329; Indels 177; Gaps 27;

Qy 6 GP-GPTDTERTPP-----ADLSAQGLEASAAKSAQAQRI--AGAEAK 46
```



```

Db 772 ---TAVTQALNELLOHVK-----AHATGAGPAGR-----YDQATDILTITVTENIFSSMGDAGE 821
QY 116 TAYDTFTSTSLADTQAAALSVLQDQAVTNKIKOTAATDEETATAAEWETKNADAVK----- 169
Db 922 MYQARIILQAQTSIDLVAIKADAEGESLENSRKLLSAAKILADATAKWVEAKGAAAHHP 881
QY 170 -----VGAQTEILAKYASDNOALDLSLGKLTISFDLIQAA--LIQSVANNKKAEL 217
Db 882 DSEEQQRLREAAEGRLMATNAQAQNAIKKLVQRLEHAQAQAASATQTIAAQAHAASA 941
QY 218 LKE-----MODNFVVGKTPATAQSLVDTQDTATATQIEKDGNAIRDAYFAGQN----- 265
Db 942 PKASAGPOLLVQSCKVAEQIPLIVQV-----RGSOAQDPSQAQALIAAQSFLQ 995
QY 266 -----ASGAVENAKNSNISINISDAKAALATATQTAETAEAKKFPDS--PILOEASQM 315
Db 996 PGKQVAAAKASVPTIQOASAWQLSQCAKNGLTALAEILRTAAQAQACGPLEVDSALS 1055
QY 316 VIQ-AEKOLKNTKPA--DGSVDPNPGTT-----VGGSKQGGSSIGSIRVSMMLDDAEN 365
Db 1056 VVQNLKDLQETKAAARDGKLPFGETMEKCTQDLGNSTKAVSSAIAKLIGELIAQGNEN 1115
QY 366 E---TASILMSGFROMIHFN-----TENPDQAAQQLAQAQ-----RAAKAA-- 406
Db 1116 YAGIAARDVAGRLSLAQARGVAALUTSDPAVQAIVLDTASDVLDKASLSLIEAKKASGH 1175
QY 407 -GDDSAALADAQKALEAALKAGQ-----QQGILNALGQI--ASAAYVAGVPPA----- 455
Db 1176 PGDPESQQLAQVAKAVTQALNRCVSLPGQRDVDNALRAVGDASKRLSLDLLPSTGTF 1235
QY 456 -----AASSIGSSVKQLYKTSKST-----GSDYKQISAGYDAYKSINDA 495
Db 1236 QBAQSRLENAAGLNAQAATELVQASRGTPQDLARASGRFGQDFSTFLEAGV-----EM 1288
QY 496 YGEARNDATP-DVINN-----VSTPALTRSVPRATEARGPEKTDQ--ALARVISGNSRT 547
Db 1289 AGQAFSQEDRAQVNVNKGISNSSSKLLLAALKALSTDPAFNLKSLQALAAARVDTDSINQ 1348
QY 548 LGDVYSQ-----VSALQSVQHIQSNPQANNE-----EIRQKLTSAVT 595
Db 1349 LITMCTQAPGQKECDNALRQLETVRELLENVPQINDMSYFGCLDSVMNSKVLGEAMT 1408
QY 596 -----KPPQFGYPVQLNSDSTQKFIKLESIFABGSTAAEIKALSFEITSFLIQ 636
Db 1409 GISQNAKGNLPEFGDAIATAS-----KALCGFTFAAAQAAYLVGVSD-PPRSQAGQ 1458
QY 637 QVILV 640
Db 1459 QGLV 1462

RESULT 7
HRPK_PSESY
ID_HRPK_PSESY STANDARD; PRT; 641 AA.
AC P41501;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Pathogenicity locus protein hrpK.
GN HRPK.
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94148760; PubMed=8106313;
RA Xiao Y., Heu S., Yi J., Lu Y., Hutcheson S.W.;
RT "Identification of a putative alternate sigma factor and
RT characterization of a multicomponent regulatory cascade controlling
RT the expression of Pseudomonas syringae pv. syringae Pss61 hrp and

```

```

hrmA genes.";
RL J. Bacteriol. 176:1025-1036(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94100578; PubMed=8274770;
RA Heu S., Hutcheson S.W.;
RT "Nucleotide sequence and properties of the hrmA locus associated with
RT the Pseudomonas syringae pv. syringae 61 hrp gene cluster.";
RL Mol. Plant Microbe Interact. 6:553-564(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03855; AAA17653.1; -.
SQ SEQUENCE 641 AA; 67678 MW; EC098941B5B46C8E CRC64;

Query Match 5.6%; Score 178; DB 1; Length 641;
Best Local Similarity 20.2%; Pred. No. 0.088;
Matches 125; Conservative 72; Mismatches 244; Indels 178; Gaps 20;

QY 53 DSVRWSILRSVAVNALMSLADKLGIASNSSSSTSRSDVDS-----TTATPTPPPT 106
Db 53 DSPSFSFGTEQSGSLMSL-----LTRSSSEST'S-SVDQSDQVSPMTSVSSTASAPT 106
QY 107 FDDYKTOQATYDITFTSTSLADIQAALVSLQDAVTNI---KOTATDEETATAAEWET 162
Db 107 AASNPNANPSTDAAFDLNSEYSPALKRWPEVAVNLPPEREQAQAKELNRPITAAWMA 166
QY 163 KNADAVKVGQITELAKYASDNOALDLSL-----GKLTSPDLQALLOSANNV 212
Db 167 R-----ENGPNAEKAMAFINANPALKYADVGVKDGNGADGKITNKOLKAFKAKNMEKADN 221
QY 213 KAEALLKEMQDN-----PVVPGKTPAIAGSLVDQTDATA----- 246
Db 222 ADKDVAKYTMEDNPGADPOSLEWRSAAVWRANMPLATAADPHHAVGAADKTDVDCNVAE 281
QY 247 -----TQIEDGNAIR-----DAYFAGQNAS-- 267
Db 282 GLKALIKSNPGLSGTLKOSSNNWSOAGFLSOVDEAGLTGRKKAHSPQVFDASNNSEWI 341
QY 268 -----GAVENAKNSNISINISDAKAALATATAKTQIAEAKKFPDSP-----IL 309
Db 342 RKSAPKNGGQFASMLSDAATLNSVAGID-----ISKLNQAVPEKPKAYTGAQKAAVMIKL 396
QY 310 QBAEQMVIAQSKDLKNIKPADGSDVPNPGTIVGSGKQGGSSIGSIRVSMMLDDAENETAS 369
Db 397 QQTQOQSVIA-----GRDLRNTEKTEAGLINE-----RIAQL--QADPDVQE 434
QY 370 ILMGFRQMIHMENTENPDQAAQCELAAQAARAKAAGDDSAALADADAKALE----- 423
Db 435 YLKNISIPQERSILRSVSDALQKATEQJQNVNSKALQTD-----LATADAKVKNHNDPP 489
QY 424 -----AALGRAGQOQGGILNALGQISAAVVSAGVPPAAASSI-----GSSVKQL 467
Db 490 DYSGAITGLSAQLQLOKDLFPDAQVPTAQVFNQPDQVQTKIADSVYVRNFSEGGALKQL 549
QY 468 YKTSKSTGSDYKTOISAGYDAYKSI--NDAYGRANDATRDVINNVSTPALTRSVPRART 525
Db 550 LGQKSDAGESLQTDADNQAAYESVLPADFVNGRESYVSTASTLQNSKKGRKILLEGKT 609
QY 526 EARG-----PEKTDQA 536
Db 610 DEEGALRWSPSSPSKVSFA 628

RESULT 8
N214_HUMAN

```



```

OS Halobacterium salinarium.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
CX NCBI_taxID=64091, 2242;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorson V., Sirogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Lechauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN
RP SEQUENCE FROM N.A.
RC SPECIES=H. salinarium; STRAIN=S9;
RX MEDLINE=96275896; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
RA Oesterheld D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
transduction.
CC
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC
CC -!- SIMILARITY: Contains 1 HAMP domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE005080; AAG19985.1; -.
CC EMBL; X95589; CAA64842.1; -.
CC PIR; E84327; E84327.
CC PIR; T46811; T46811.
CC HSSP; P02942; 1QU7.
CC InterPro; IPR004089; Chmtaxis_transd.
CC InterPro; IPR003660; HAMP.
CC InterPro; IPR004090; Me_chemotaxis.
CC Pfam; PF00015; MCPsignal; 1.
CC PRINTS; PR00260; CHEMTRNSDUCR.
CC SMART; SM00304; HAMP; 1.
CC SMART; SM00283; HA; 1.
CC PROSITE; PS01111; CHEMOTAXIS_TRANSDUC_2; 1.
CC PROSITE; PS00885; HAMP; 1.
CC Transducer; Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 90 100 POTENTIAL.
FT DOMAIN 164 217 HAMP.
FT DOMAIN 236 476 METHYL-ACCEPTING TRANSDUCER.
FT SEQUENCE 545 AA; 57070 MW; FDD670389C2F428B CRC64;
SQ
Query Match 5.5%; Score 176.5; DB 1; Length 545;
Best Local Similarity 21.4%; Pred. No. 0.085;
Matches 98; Conservative 79; Mismatches 195; Indels 87; Gaps 17;
QY 12 ETERPPADLSAQGLEASAAKSAEQAQRIAGAEAKPKESKTSVERWSTLRSAVNALMSL 71
DB 118 EEERA-----EERAREKAEQQAER---QTAEASAKQARSRABIEQLAADLESQ 169
QY 72 ADKLGIASSNSSSTSRADVDSTATATPTPPPT-FDDYKTKQAQYAVDTI-----FTST 125
DB 170 ATEVG-ATLEASDGLTARVDATDNAAIEAVATVNDMLTMTERTIDEIQGFSTNVT 228

```

```

QY 126 SLADIQAALVSLQDAVTNIKDT-----AATDEE-----TAIAAEWETKNADAVKYGAQITE 176
DB 229 ASREATAGAKIEIQASQTSVESQETIAAGTDDQREQLSEVAEEMDSYATVEEVAATAQS 288
QY 177 LAKYASDNQALLDSLGKLT-----SFDLLQAALLQSVAANNKAAEILLKEMQDNPPVPGK 230
DB 289 VADTAADTTVA-TAGKQTAEDAIDAIDAVOETQMTTVANVDALEDTTTEID----- 340
QY 231 TPAIAQSLVDTQDTATATQIEKDGNAIRDVAFAGNASG-----AVENAKSNNSI 279
DB 341 IAEILSIDIAEQTNMLA-LANATIEAARAGSGSGGDFAVVADEVKELATESQSAKDI 398
QY 280 SN-IDSAAIAIYAKTOIAEQAQKFPSPILQEAQOMVIOAKDLKNIKPKADGSDVNPNG 338
DB 399 AELIEVQSQATATTVEBI-----RVAQRVNDGAAAVEETVDAFGAVTENIQ 445
QY 339 TTVGSQKQGGSIGSIKVSMLDDAENETASILMSGFRQMHMENTENPDSQAQQLAA 398
DB 446 ETTDGVQE-----ISQAMDEQAQORSERVSS-----VDDIATISQATAD 484
QY 399 QARAAGAGGDSMAAALADAKALEAALGRAGQOQGIIN 437
DB 485 RAENVSAASEEQ-TASITEVTSSLSQSLAAQADTLEDRIN 522

```

```

RESULT 10
YMW6_YEAST STANDARD; PRT; 1140 AA.
AC 004893;
DT 01-NOV-1997 (Rel. 35, Created)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 113.1 kDa protein in PRE5-FET4 intergenic region.
GN YMS317W OR YM9924.09.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_taxID=4932;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RT Nature 387:90-93(1997).
RL
CC -!- DOMAIN: Contains many Ser/Thr-rich domain and repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).
CC
CC EMBL; Z54141; CAA90835.1; -.
CC GenOnline; 142998; -.
CC SGB; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;
Query Match 5.5%; Score 174.5; DB 1; Length 1140;
Best Local Similarity 17.3%; Pred. No. 0.27; 273; Indels 167; Gaps 17;
Matches 118; Conservative 126; Mismatches 126; Indels 167; Gaps 17;
QY 79 SSNSSTSRGADVDSTATATPTPPPT-----PTFDDYKTKQAQYAVDTIFSTIS 126
DB 23 SSTSTTPTTSTTSITSTSKVTSPEIIVSSSTSLVSSVVPFTSSSLSSDIASILS 82

```



```

RP RESULT 12
RX MINT MOUSE
RA MINT MOUSE STANDARD; PRT; 3644 AA.
AC Q62504; O80TN9; Q99P84; Q9QW2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE MSX2-interacting protein (SMART/HDAC1 associated repressor protein).
GN MINT OR SHARP OR KIAA0929.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE OF 49-3644 FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR
RP LOCATION, TISSUE SPECIFICITY, DNA-BINDING, AND INTERACTION WITH MSX2.
RC TISSUE=Testis;
RX MEDLINE=99379811; PubMed=10451362;
RA Newberry E.P., Latifi T., Towler D.A.;
RT "The RRM domain of MINT, a novel msx2 binding protein, recognizes and
RT regulates the rat osteocalcin promoter.";
RL Biochemistry 38:10678-10690(1999).
[2]
RP SEQUENCE OF 1-112 FROM N.A.
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
RA Hashizume W., Hayashida K., Hirozane T., Hori F., Imotani K.,
RA Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y., Kondo S., Konno H.,
RA Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
RA Ohno M., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki D.,
RA Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki K.,
RA Watabiki A., Muramatsu M., Hayashizaki Y.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
RP PHE-762; PHE-773 AND LEU-933.
RC STRAIN=ICR; TISSUE=Brain;
RA Sakamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaichi M.;
RT "MINT/spen negatively regulates Notch signaling by inhibiting RBP-
RT J/Su(H) activity.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE OF 318-578 FROM N.A.
RC TISSUE=Cochlea;
RX MEDLINE=97237053; PubMed=9119401;
RA Crozet F., El-Amraoui A., Blanchard S., Lenoir M., Ripoll C., Vago P.,
RA Hamel C., Fzames C., Levi-Acobas F., Depetris D., Mattei M.-G.,
RA Weil D., Pujol R., Petit C.;
RT "Cloning of the genes encoding two murine and human cochlear
RT unconventional type I myosins.";
RL Genomics 40:332-341(1997).
[5]
RP SEQUENCE OF 2598-3644 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=2257291; PubMed=12693553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT I. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
[6]
RP TISSUE SPECIFICITY.
RX MEDLINE=22261914; PubMed=12374742;
RA Oswald F., Kostecka U., Astrahantseff K., Bourteelle S., Dillinger K.,
RA Zechner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Lipray S.,
RA Schmid R.M.;
RT "SHARP is a novel component of the Notch/RBP-Jkappa signalling
RT pathway.";
RL EMBO J. 21:5417-5426(2002).
[7]
RP

```

```

RP FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=22483652; PubMed=12594956;
RA Kuroda K., Han H., Tani S., Tanigaki K., Tun T., Furukawa T.,
RA Taniguchi Y., Kurooka H., Hamada Y., Toyokuni S., Honjo T.;
RT "Regulation of marginal zone B cell development by MINT, a suppressor
RT of Notch/RBP-J signaling pathway.";
RL Immunity 18:301-312(2003).
CC -!- FUNCTION: Essential corepressor protein, which probably regulates
CC different key pathways such as the Notch pathway. Negative
CC regulator of the Notch pathway via its interaction with RBPSUH,
CC which prevents the association between NOTCH1 and RBPSUH, and
CC therefore suppresses the transactivation activity of Notch
CC signaling. Blocks the differentiation of precursor B cells into
CC marginal zone B cells. Probably represses transcription via the
CC recruitment of large complexes containing histone deacetylase
CC proteins. May bind both to DNA and RNA.
CC -!- SUBUNIT: Interacts with NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and
CC MYALL1. Interacts with the nuclear receptors RAR and PPARδ.
CC Interacts with RAR in absence of ligand. Bind to the steroid
CC receptor RNA coactivator SRA (by similarity). Interacts with MSX2.
CC Interacts with RBPSUH; this interaction may prevent the
CC interaction between RBPSUH and NOTCH1.
CC -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q62504-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q62504-2; Sequence=VSP_008564;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower
CC level in brain, lung, spleen, liver and kidney. Weakly expressed
CC in cardiac and skeletal muscles and ovary. In spleen, it is
CC expressed in follicular B-cells, while it is weakly expressed in
CC marginal zone B-cells.
CC -!- DOMAIN: The RID domain mediates the interaction with nuclear
CC receptors.
CC -!- DOMAIN: The SPOC domain, which mediates the interaction with
CC NCOR2, is essential for the repressive activity (By similarity).
CC -!- SIMILARITY: Belongs to the Spen family.
CC -!- SIMILARITY: Contains 1 RID (receptor interacting) domain.
CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC -!- SIMILARITY: Contains 1 SPOC domain.
CC -!- CAUTION: Ref.4 sequence differs from that shown due to multiple
CC frameshifts and conflicts that create stop codons.
CC -!- CAUTION: Ref.5 sequence differs from that shown due to what seems
CC to be the presence of intronic sequence in the cDNA.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; BY26481; -; NOT ANNOTATED CDS.
DR EMBL; AF156529; AAD55931.1; ALT_INIT.
DR EMBL; AB055980; BAB32786.1; -.
DR EMBL; Z78160; CAB01562.1; ALT_SEQ.
DR EMBL; AK122402; BAC65684.2; ALT_SEQ.
DR MGD; MGI:1991706; Mnt.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 3.
DR SMART; SM00360; RRM; 3.
DR PROSITE; PS0102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE; PS00917; SPOC; 1.
KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
KW RNA-binding; Repeat; Coiled coil; Alternative splicing; Polymorphism.
FT DOMAIN 1 574 DNA-BINDING.
FT DOMAIN 6 81 RNA-BINDING (RRM) 1.
FT DOMAIN 336 416 RNA-BINDING (RRM) 2.

```

```

FT DOMAIN 439 514 RNA-BINDING (RRM) 3.
FT DOMAIN 518 590 RNA-BINDING (RRM) 4.
FT DOMAIN 559 575 COILED COIL (POTENTIAL).
FT DOMAIN 822 850 COILED COIL (POTENTIAL).
FT DOMAIN 1185 1206 COILED COIL (POTENTIAL).
FT DOMAIN 1509 1544 COILED COIL (POTENTIAL).
FT DOMAIN 1607 1627 COILED COIL (POTENTIAL).
FT DOMAIN 2216 2704 RID.
FT DOMAIN 3478 3644 SPOC.
FT DOMAIN 2138 2462 INTERACTION WITH MSX2.
FT DOMAIN 2706 2845 INTERACTION WITH RBPSUH.
FT DOMAIN 125 277 ARG-RICH.
FT DOMAIN 236 326 SER-RICH.
FT DOMAIN 648 721 TYR-RICH.
FT DOMAIN 702 832 ARG-RICH.
FT DOMAIN 2101 2233 ALA-RICH.
FT DOMAIN 2377 2518 PRO-RICH.
FT DOMAIN 2950 3475 PRO-RICH.
FT VARSPLIC 618 640 Missing (in isoform 2).
/FTid=vsp_008564.

FT VARIANT 348 348 I -> T.
FT VARIANT 762 762 S -> F.
FT VARIANT 773 773 S -> F.
FT VARIANT 933 933 S -> L.
FT CONFLICT 754 754 R -> G (IN REF. 3).
FT CONFLICT 1524 1524 D -> A (IN REF. 3).
FT CONFLICT 1560 1560 H -> Y (IN REF. 3).
FT CONFLICT 1570 1570 F -> L (IN REF. 3).
FT CONFLICT 1574 1574 R -> G (IN REF. 3).
FT CONFLICT 1609 1609 Q -> R (IN REF. 3).
FT CONFLICT 1659 1659 I -> V (IN REF. 3).
FT CONFLICT 1669 1669 S -> F (IN REF. 3).
FT CONFLICT 1705 1705 V -> F (IN REF. 3).
FT CONFLICT 1815 1815 A -> V (IN REF. 3).
FT CONFLICT 2097 2097 G -> A (IN REF. 3).
FT CONFLICT 2201 2202 MISSING (IN REF. 3).
FT CONFLICT 2322 2322 A -> V (IN REF. 3).
FT CONFLICT 2385 2385 P -> Q (IN REF. 3).
FT CONFLICT 2502 2502 R -> K (IN REF. 3).
FT CONFLICT 2505 2505 E -> K (IN REF. 3).
FT CONFLICT 2519 2519 D -> N (IN REF. 3).
FT CONFLICT 2554 2554 T -> S (IN REF. 3).
FT CONFLICT 2679 2688 LVSTPAGPVN -> VGEHPWARD (IN REF. 3).
FT CONFLICT 3010 3010 L -> P (IN REF. 3 AND 5).
SQ SEQUENCE 3644 AA; 398750 MW; 9C7EC49A81A7DA4A CRC64;

Query Match
Best Local Similarity 5.4%; Score 172.5; DB 1; Length 3644;
Matches 130; Conservative 111; Mismatches 314; Indels 143; Gaps 19;

QY 24 QGLEASANKSAEQRITAGAA-KPKESKTSVERWSIL-RSAYNALMSLADKILGIASSN 81
Db 1559 QHLERKSESDFPFPGRLYGRQASEGANSSTDSVQEPVLFHSRPFELTRMOCK-----EK 1613
QY 82 SSSSTSRADVDSITATPTPPPPFDYKQACTAYTITSTSLADIQAALVSLQDAV 141
Db 1614 EKQKPKAEKQEBETHPKTPEPAETKEPEPK-----APVSAGLPAVITV 1661
QY 142 TNIKDTATDETAIAAEWETKNADAVKVGQITELAKYASDNOAI-----L 188
Db 1662 VTPEPASSAPEKAEEAAEAPSPAGEKPAEPAPVSEETKLVSPEVSPVPEQPSDVPGE 1721
QY 189 DSLGKLTFDLLQALLOSVANNKAEALLKEMQNPVVPKTPAIAQLVD-----240
Db 1722 DSRDSQDS-----AALAPSAPQESAATDAVPCWNAEPLTPGTTSVQSVSSVDPKPSQP 1776
QY 241 -----QTDATATOIEK-----DGNALRDVAGQNASGAVENAKSNNSINISAK 286
Db 1777 LSKLTQRSEAEKGEKPEPTTPSTPDATQNAVASEAQPPASDVEANPPVAAKDRKT 1836
QY 287 AAIATAKTQIAEAKKFPDPSILQAE-----QMWIQAEKDLKN 325
Db 1837 NKSKESKTSVQAAASVVEKPVTRKSERIDREKLKSSPRGEAQKLELXWAEKIRTI 1896

```

```

QY 326 IKPADGSDVNPCTTVGGSKQSGSIGSRVSNMLDDAENE-----TASILMSGF 375
Db 1897 ASKSGGDTHEHPSPPLSRRRNRVSVYATWTDHESRSPAKEPVEQPRVTKRLREL 1956
QY 376 RQMIHMFNT-----ENPDSQAQOELAAQABAKA-----AGDSSAALADAQ 419
Db 1957 QEAVVPTTFRGRPEKTRRAEDGEHERKEPAETPRFAEGRWSRSPKSAANA-----2011
QY 420 KALEAALGAGQOQILNALGOIASAAVVSAGVPPA-----AASSIGSVKQLYKTSKSTG 475
Db 2012 -----GPGQKRGNEQVEAAEAGAQAQSTREGNPKSGRGEAASEPKRDRDSTDKSGP 2067
QY 476 SDYKTIQISAGYDAYKSIINDAYGEARNDATRDVINNV-----STPALRSVPRARTEARGPE 531
Db 2068 DTFPEVLEKRPPEKTYKSGRGRAR--STRSGMDRAAHQSRLEMAAAGAAQADKEAGP-2124
QY 532 KTDQALARVISGNSRTLGDVYSQVSAIQSMQIIOQNPOANNEEIRQKLTSAVTKPQFG 591
Db 2125 -----AAAFQSESPQKSGSSPQLA-----NNPADPDREAEESASASTAPPEGT 2171
QY 592 YPVVQLSNDSTOKFIACL-----ESLFAEGRSRTAAEIK 625
Db 2172 QLAHQIIELEGAVQNIAKLPEPSAAASKGTATATATA 2209

RESULT 13
SPAA_STRDO STANDARD; PRT; 1528 AA.
AC P21979;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell surface antigen I/II precursor.
GN SPAA.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91310320; PubMed=1855987;
RA Lapollia R.J., Haron J.A., Kelly C.G., Taylor W.R., Bohart C.,
RA Hendricks M., Pyatt J., Graff R.T., Ma J.K.-C., Lehner T.;
RT "Sequence and structural analysis of surface protein antigen I/II
RT (Spaa) of Streptococcus sobrinus."
RL Infect. Immun. 59:2677-2685(1991).
RN [2]
RP SEQUENCE OF 423-817 FROM N.A.
RX MEDLINE=90299827; PubMed=1694526;
RA Goldschmidt R.M., Thoren-Gordon M., Curtiss R. III;
RT "Regions of the Streptococcus sobrinus spaa gene encoding major
RT determinants of antigen I."
RL J. Bacteriol. 172:3988-4001(1990).
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC -!- MISCELLANEOUS: IMMUNODOMINANT DETERMINANTS ARE LOCATED IN THE
CC C-TERMINAL TWO-THIRDS OF THE SPAA PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SPAP/SSP5/SPAA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X57841; CAA40973.1; -.
DR EMBL; M38210; AAA26977.1; -.
DR InterPro; IPR001899; Gram_pos_anchor.
DR Pfam; PF00746; Gram_pos_anchor; 1.

```



```

Db 594 DNAKHTSYGSSGASLDSRTTTSISVSSNTT-----QLVSTCTSESYS----- 640
Qy 394 QELAAQARAAGDGAALADQAALAKAGAGQOQILNALGQIASAAVVSAGVP 453
Db 641 -----DPSFPAISATATTT-----ESNLTWT-----ITASCSTDSNFP 673
Qy 454 PAAASS-----IGSSVKQL-----YKTSKSTGSDYKTKQISAGYDAYKSI-----ND 494
Db 674 TSAASSTDETAFTRTISTSCSTLNGASTQSELITSPMKNTVTVVPASSFPSTTTTCLEND 733
Qy 495 --AYGRARNDATRVINNVSTPALTRSPVPRARTARGPEKTDQALARVISGNSRTLGDVY 552
Db 734 DTAESSIYTEVNAATIIN---PGETSSLASDPATSEKNEPSTSVKSTSNEGTSST----- 785
Qy 553 SQVSALQSVQIIQOSNPOANNEEUR-----QKLSAVTKPPQFGYPVYVQLSND 600
Db 786 --TTYQOTVATLVAKPSSTSLGARTTGSNGRSTTSQDGSAMHQPTSSYITQKEGTS 843
Qy 601 STQKPIAKLESFPAEGSRTAAEIKALGFETNSL-----FQQVL 639
Db 844 TTKALSA-----YEGATPLSI-----PQCNSLAGTIAAFVAVL 878

```

RESULT 15

```

HTR4 HALSA
ID HTR4 HALSA STANDARD; PRT; 810 AA.
AC Q48317;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Halobacterial transducer protein IV.
GN HTR5 OR HTRIV.
OS Halobacterium salinarum.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_taxID=2242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S9;
RX MEDLINE=96275896; PubMed=8674984;
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,
RA Oesterhelt D.;
RT "A family of halobacterial transducer proteins.";
RL FEMS Microbiol. Lett. 139:161-168(1996).
CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal
CC transduction.
CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -!- SIMILARITY: Contains 2 HAMP domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X95589; CAA64841.1; -.
CC PIR; T46810; T46810.
CC HSSP; P02942; 10U7.
CC InterPro; IPR004089; Chmtaxis_transd.
CC InterPro; IPR003660; HAMP.
CC Pfam; PF00672; HAMP; 1.
CC Pfam; PF00015; MCPsignal; 1.
CC SMART; SM00304; HAMP; 2.
CC SMART; SM00283; MA; 1.
CC PROSITE; PS50111; CHEMOTAXIS_TRANSDCUC_2; 1.
CC PROSITE; PS50885; HAMP; 2.
KW Transducer; Transmembrane; Repeat.
FT DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 POTENTIAL.
FT DOMAIN 60 323 EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 324 344 POTENTIAL.
FT DOMAIN 345 810 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 345 397 HAMP 1.
FT DOMAIN 439 493 HAMP 2.
FT DOMAIN 512 748 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 810 AA; 85207 MW; 37B0F6046A39D9BA CRC84;

Query Match
Best Local Similarity 19.7%; Pred. No. 0.36;
Matches 144; Conservative 103; Mismatches 316; Indels 168; Gaps 26;

Qy 11 DETERPPADLSAQ-----GLEASANKSA-----EAORIA-----GAEA 45
Db 86 DWLERNEQSVLIASNNPRLGNTTAADKQAYVTVQLVAAELDADRIADVHVADPTVGGASD 145
Qy 46 KPRESKTDSEVERSIILRSATNALMSLADKLG-----LIASSNS-----SSSTSRSDVD 93
Db 146 ARIVASTDEDRGTRVSADTHPWDRTRSGRTVTSNPNYRTAGGQVVSSMSVADLT 205
Qy 94 STTATAPTPTPTDDYKTKQATAYDTIFTSTSLADIQAALVSLQDAVTNI----- 144
Db 206 HVLVVEYTA-----GDLSDQFGAGIDGTFTQVVVRPTSDATAVLFSDDAGTDAVGQYIPDR 260
Qy 145 --KDTATDEETALAAEWETKNADAV-----KVGQAQITELAK-----YASDNQ 185
Db 261 SQSEI PALDSATEQGGTPTTKDSVLDRYVAAAYTTPGKNVWVVKHAPSESASFALSNQ 320
Qy 186 AILDSLG-----KLTSPDLL-----QAALLQSVANNKAAELKEMQDNPVVPKTPAI 234
Db 321 IRGILGFIILVGVVVGVTIGRTAAAVQSL--SAAAEIEAGNYDNDVASSRDEI 378
Qy 235 AQ---SLVDQTDATATQIEKGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKA-AIA 290
Db 379 GOLFSAGISMRDALVTQIDAEAAAREQATEAQQDAE--ABERAEADAREAEADAKADAEA 436
Qy 291 TAKTQIAEAKKXFPDPILOEAEQVIOAEKDLKNIKPADGSDVPNPCTVGGSKQOGSS 350
Db 437 LAELEQAQER-----YSDVMAACADGDUUTRMPA----- 466
Qy 351 IGSTRVSMILLDDAENETASILNSGFRQMIHMENTENPDSSQAQQLAAQARAARAAAGDSDS 410
Db 467 -----DDTDNEAMAAIASFNEMLAQWHEHTIIDIQEFADAVATASEAEVGNAD- 515
Qy 411 AAAALADAQKALEALGKAGQOQILNAL-GQIA--SAAVVSAGVPPAAASSIGSSVKQL 467
Db 516 AERASGVSESVQEIAGAAGEQRMMLDTVSGEMTDLSSAIEEVAAGSADSVAEHSHQTAEI 575
Qy 468 YKTSKSTGSD-----YKTIQISAGYDAYKSIINDAYGRAEN--DATRDVINNVSTPALTR 518
Db 576 ARGEQTAEDAIERSISVQBAIDATVQNVFALDDQWAEISEIVDLISDIAEQTNMLALNA 635
Qy 519 SVPRARTEARG-----PKTDQALARVISGNSRTILGDYVSQVSALQSVQMII 565
Db 636 NIEAARADKSGDGFVAVVADSVKDLABETQESAGDI---ERRITEVQSQTATVAERAA 691
Qy 566 QSNPQAN---NEIROKLSAVTKPPQFGYPVYVQLSN-----DSTQKFAKLESFPAEG 616
Db 692 EESMDAGIDAVEEVVDFAVSDHADETTGTGVEIISDTTDDQASSTEEAVSMTEEVADLS 751
Qy 617 SRTAAEIKALS 627
Db 752 DSTAGEAQSVS 762

```

Search completed: March 24, 2004, 05:56:56
Job time : 28.9698 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:47:55 ; Search time 49.0374 Seconds
(without alignments)
4188.690 Million cell updates/sec

Title: US-10-608-559-2

Perfect score: 3187

Sequence: 1 MNVIGPDIETERTPAD.....SLFIQQVLVNTGSLYSGVLQ 651

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3187	100.0	651	16 Q9Z7H7	Q9Z7H7 chlamydia p
2	2205	69.2	715	2 Q46166	Q46166 chlamydia p
3	1276.5	40.1	636	16 Q824X3	Q824X3 chlamydia p
4	899.5	28.2	647	16 Q84627	Q84627 chlamydia t
5	247	7.8	2310	16 Q8CMU7	Q8CMU7 staphylococ
6	232.5	7.3	2055	2 Q85472	Q85472 abiotrophia
7	232	7.3	1822	2 Q07290	Q07290 streptococ
8	231.5	7.3	2478	2 Q9RL69	Q9RL69 staphylococ
9	229.5	7.2	2478	2 Q9LCH2	Q9LCH2 staphylococ
10	223.5	7.0	2481	16 Q99QK6	Q99QK6 staphylococ
11	221.5	7.0	973	16 Q8XQJ4	Q8XQJ4 escherichia
12	220	6.9	6713	16 Q99U54	Q99U54 staphylococ
13	217.5	6.8	6713	16 Q93IR6	Q93IR6 staphylococ
14	215.5	6.8	2016	5 Q9BI70	Q9BI70 pleocitreure
15	214	6.7	842	5 Q96RNL	Q96RNL plasmodium
16	214	6.7	2275	16 Q8NUJ3	Q8NUJ3 staphylococ

17	214	6.7	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
18	213.5	6.7	3360	16 Q88XB6	Q88XB6 lactobacilli
19	212	6.7	855	5 Q86DH2	Q86DH2 plasmodium
20	211	6.6	9439	16 Q8CP76	Q8CP76 staphylococ
21	210.5	6.6	1795	16 Q91CJ9	Q91CJ9 staphylococ
22	210	6.6	2271	16 Q99QI4	Q99QI4 staphylococ
23	209	6.6	625	5 Q86RN6	Q86RN6 plasmodium
24	209	6.6	2167	2 Q845L8	Q845L8 mycoplasma
25	208.5	6.5	3072	2 Q939N5	Q939N5 streptococ
26	208.5	6.5	4776	16 Q97P71	Q97P71 streptococ
27	207.5	6.5	856	5 Q86RM6	Q86RM6 plasmodium
28	207	6.5	621	5 Q86RM2	Q86RM2 plasmodium
29	206.5	6.5	762	16 Q89IU6	Q89IU6 bradyrhizob
30	205	6.4	625	5 Q86RN2	Q86RN2 plasmodium
31	204	6.4	803	5 Q86RN4	Q86RN4 plasmodium
32	204	6.4	1313	5 Q9NE28	Q9NE28 leishmania
33	204	6.4	3381	2 Q9KX33	Q9KX33 streptococ
34	201.5	6.3	849	5 Q86R71	Q86R71 plasmodium
35	201	6.3	1336	16 Q9ACV2	Q9ACV2 streptomyce
36	200.5	6.3	1545	16 Q9RDQ1	Q9RDQ1 streptomyce
37	200	6.3	732	5 Q86RM7	Q86RM7 plasmodium
38	200	6.3	842	5 Q86RN3	Q86RN3 plasmodium
39	199	6.2	1365	2 Q49525	Q49525 mycoplasma
40	199	6.2	2178	2 Q9KWR3	Q9KWR3 streptococ
41	198.5	6.2	1932	16 Q7UQ94	Q7UQ94 rhodospirell
42	198	6.2	849	5 Q86RM3	Q86RM3 plasmodium
43	198	6.2	1222	16 Q87XX8	Q87XX8 pseudomonas
44	197.5	6.2	838	5 Q86RN7	Q86RN7 plasmodium
45	197.5	6.2	2283	2 Q8VQ99	Q8VQ99 staphylococ

ALIGNMENTS

RESULT 1

Q9Z7H7
ID Q9Z7H7 PRELIMINARY; PRT; 651 AA.
AC Q9Z7H7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CHLPN 76 kDa HOMOLOG 1 (C1622) (Hypothetical protein).
GN CPN0728 OR CPJ0728 OR CP0018 OR CPB0756.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CNL029;
RX MEDLINE=9206606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RA "Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*."; Nat. Genet. 21:385-386 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., Bass S.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Kouri H., Craven B., Bowman C., Dodson R.,
RA Winn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RA "Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39."; Nucleic Acids Res. 28:1397-1406 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RA "Comparison of whole genome sequences of *Chlamydia pneumoniae* J138

```
RT from Japan and CWL029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-TW-183;
RA Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001654; AAD18867.1; -
DR ENBL; AE002165; AAR37914.1; -
DR ENBL; AP002547; BAA98935.1; -
DR EMBL; AE017159; AAP98685.1; -
DR PIR; D72042; D72042.
DR PIR; E86581; E86581.
DR PHCI-2DPAGE; Q927H7; -.
DR TIGR; CP0018; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 651 AA; 68217 MW; 47AB6C3FF2FF0123 CRC64;

Query Match 100.0%; Score 3187; DB 16; Length 651;
Best Local Similarity 100.0%; Pred. No. 2.9e-137;
Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAAQRIAGAEAKPKESKTSVERWSI 60
Db 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAAQRIAGAEAKPKESKTSVERWSI 60
QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPPPTTDDYKTKQAQTAYDT 120
Db 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPPPTTDDYKTKQAQTAYDT 120
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 180
Db 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 180
QY 181 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKAAELLKEMODNPVVPKTPAIAQSLVD 240
Db 181 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKAAELLKEMODNPVVPKTPAIAQSLVD 240
QY 241 QTDATATQIEKGNAINRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 241 QTDATATQIEKGNAINRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
QY 301 KKPDPSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSIRVSMML 360
Db 301 KKPDPSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSIRVSMML 360
QY 361 DDAENETASILMSGFQMIHMENTENPDQAQOELAAQARAACAAGDSDSAAALADAQK 420
Db 361 DDAENETASILMSGFQMIHMENTENPDQAQOELAAQARAACAAGDSDSAAALADAQK 420
QY 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVPPAAASSIGSSVKLYKTSKSTGSDYKT 480
Db 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVPPAAASSIGSSVKLYKTSKSTGSDYKT 480
QY 481 QISAGYDAYKINDAYGRANDATRVINNVSTPALTRSPVPRATEARGPEKTDQALARY 540
Db 481 QISAGYDAYKINDAYGRANDATRVINNVSTPALTRSPVPRATEARGPEKTDQALARY 540
QY 541 ISGNSRTLGDVYSQVSALQSVQMLIQSNPQANNEIEIKLTSAVTKPPQFGYPVQLSND 600
Db 541 ISGNSRTLGDVYSQVSALQSVQMLIQSNPQANNEIEIKLTSAVTKPPQFGYPVQLSND 600
QY 601 STOKFIAKLSLFAEGSRTAAEIKALSFETNSLFIQQLVNLIGSLYSGYLQ 651
Db 601 STOKFIAKLSLFAEGSRTAAEIKALSFETNSLFIQQLVNLIGSLYSGYLQ 651

RESULT 2
Q46166
ID Q46166 PRELIMINARY; PRT; 715 AA.

AC Q46166;
DT 01-NOV-1996 (TrEMBLrel. 01, Created);
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update);
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update);
DE 76 kDa protein.
OS Chlamydia pneumoniae (Chlamydiales; Chlamydiaceae; Chlamydia).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94156481; PubMed=7509320;
RA Perez-Melgosa M., Kuo C.-C., Campbell L.;
RT "Isolation and characterization of a gene encoding a Chlamydia
RT pneumoniae 76-kilodalton protein containing a species-specific
RT epitope.";
RL Infect. Immun. 62:880-886(1994).
DR EMBL; L23921; AAA23117.1; -.
DR PIR; I40729; I40729.
SQ SEQUENCE 715 AA; 76626 MW; 3F01C9A59DE964C5 CRC64;

Query Match 69.2%; Score 2205; DB 2; Length 715;
Best Local Similarity 99.6%; Pred. No. 1.5e-92;
Matches 452; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNPPIGPGPIDETERTPPADLSAQGLEASAAKSAEAAQRIAGAEAKPKESKTSVERWSI 60
Db 257 LVNPIGPGPIDETERTPPADLSAQGLEASAAKSAEAAQRIAGAEAKPKESKTSVERWSI 316
QY 61 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPPPTTDDYKTKQAQTAYDT 120
Db 317 LRSVAVNMLSLADKLGIASSNSSSTSRSDVDSSTTATATPPPTTDDYKTKQAQTAYDT 376
QY 121 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 180
Db 377 IFTSTSLADIQAALVSLQDAVTNIKDTAATDEETAIAAEWETKNADAVKVGQITELAKY 436
QY 181 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKAAELLKEMODNPVVPKTPAIAQSLVD 240
Db 437 ASDNQAILDSLGKLTSPDLLQAALLQSVANNKAAELLKEMODNPVVPKTPAIAQSLVD 496
QY 241 QTDATATQIEKGNAINRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 497 QTDATATQIEKGNAINRDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 556
QY 301 KKPDPSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSIRVSMML 360
Db 557 KKPDPSPILQEAQMVIOAEKDLKNIKPADGSDVPNPPTTGGSKQGGSSIGSIRVSMML 616
QY 361 DDAENETASILMSGFQMIHMENTENPDQAQOELAAQARAACAAGDSDSAAALADAQK 420
Db 617 DDAENETASILMSGFQMIHMENTENPDQAQOELAAQARAACAAGDSDSAAALADAQK 676
QY 421 ALEAALGKAGQGGQILNALGQIASAAVVSAGVPP 454
Db 677 ALEAALGKAGQGGQILNALGQIASAAVVSAGVLP 710

RESULT 3
Q824X3
ID Q824X3 PRELIMINARY; PRT; 636 AA.
AC Q824X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created);
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update);
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update);
DE Hypothetical protein.
GN CCA00015.
OS Chlamydia caviae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiales; Chlamydiales.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364;
```


RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J., Haft D.H., Khouri B., Federova E., Carthy H.A.,
RA Umayam L.A., Holt D.H., Peterson J., Beanan M.J., White O.,
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavov P.M.,
RA Fraser C.M.;
RT "Genome sequence of Chlamydomonas reinhardtii"
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae";
RL Nucleic Acids Res. 31:2134-2147(2003).
RL EMBL; AEO16994; AAP04768.1; --
DR TIGR; CC000015; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 636 AA; 67020 MW; C734944DCE8CA302 CRC64;

Query Match 40.1%; Score 1276.5; DB 16; Length 636;
Best Local Similarity 43.5%; Pred. No. 2.2e-50;
Matches 288; Conservative 118; Mismatches 219; Indels 37; Gaps 10

Qy	1	MVNIPIGPGPIDETERTPPADISLAOGLEASANKSABAGRIAGAKPKESK--TDSVERW 58
Dd	1	MVNEFV--GPIDESKNIPADI-STLGMQASAAARRSEAOISITGIAGKSGSQSPSVETVGR 58
Qy	59	SILSESVAALMSLDKLGLIASNSSSSPSRSADVSTTATPTPPPFPFDYDKYCAQTAY 118
Dd	59	SFLSARSKLASLFDKI---SFFSGKTT-----PQTFDEAKTQAESA 99
Qy	119	DTIFTSLADIQAALYSLODAVTNIIKDTAATDBETAIAAEWTKNADAVVGACITELA 178
Dd	100	TALQSAITYDFQKTALQQQLQDAVKQBOLATTDAEKATAVTKTALEAKSKLTTLNLQ 159
Qy	179	KYASDNQAILDSLGLKFTDLLQALQLSVANNKAABELLKEMQNPFVPGKTPAIAQSL 238
Dd	160	AALTENOKLLEAIKTTSSMDQIMGAAGGVETNKTTAEELIKQLKE---AGVSYPVIDDL 215
Qy	239	VDOTDATATQLEKDGNAIRDAYFAGONASGAVENAKSNNSINSIDSAKAATAKTQIAE 298
Dd	216	EKQTTTGTVTEADALISEAYAAGKNSTAAVGOQAQNNSPANIEASKQTIANAKRVIED 275
Qy	299	AQKFPPSPFIQEAEQMVIQAEKDLKNIKPADGSDVP--NPGT--TVGGSKQGSSSIGSI 354
Dd	276	ALKLPDSPILKAALKEQQAADLIINVKPGSGSDVIPGPGFAPGSVGTSONRGATLGEV 335
Qy	355	RVSMLDDAENETASTILMSGFRQMHPMTENPNOSQAQOBLAQARAKA--AGDDSA 411
Dd	336	RVSMLLTDVDMNETAIIIMQGFRRMIDNFHDCNQSFPLAEIIMQVTDLSQTQINPADEA 395
Qy	412	AAALADAKALEAALGRAGQOQGIILNALGQTASNAVTSAGVPPAAAASSIGSVQLKYTS 471
Dd	396	TAQCFEIQTQTDALQOTAGODGMINALGATTAAISTGTAPIASANQGGSAVKQLKYTG 455
Qy	472	KSTGS--DYKTQISAGYDAYKSINDAYGRAENDATRVDINNVSPTALTRSPVPRATEARG 529
Dd	456	STAASSKSYADSLSAGYGAVQSINDVTSRS--SASNREVLDTSTPALITQTVSRTEPRPD 514
Qy	530	PEKTDQALARVISGNSTLGDVYQVSALQSWMOIIOSNPOANNEEROKLTSAVTKPQP 589
Dd	515	NDNAAQRFARTIAANSNTLGDVASVGVLTLLGVLNQNNPOANEEDIKQLTSETVTKAPQ 574
Qy	590	FGPYVOLNSDSTOKFTAKLESFAEGSRTAETIKALSFEINSLFIQOVLINIGLSYGY 649
Dd	575	SGYPHVOLNSDSTKKFTAQLENFVQGSKRLEAKEAFKQPLFIQOVLNVNASLFSGY 634
Qy	650	LQ 651
Dd	635	LQ 636

RESULT 4
O84627 PRELIMINARY; PRT; 647 AA.
AC O84627;
Dt 01-NOV-1998 (TrEMBLrel. 08, Created)
Dt 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

[illegible]


```

Db 1798 NAVQQAADKKAAIENDPALTEEEKDAKAKADYDAEAKAKDAIDAATSNADVTAKQDAGK 1857
QY 506 DVINNV-STPA---LTRSVPRARTEARGPEKTDQALARVISGNSRTILGDVYSQVSALQS 560
Db 1858 DAINAVPQTPTAKTDAKNAVDAQADTKKSAIENDPALTR--BEKDAVKAKVDAAEKAKKD 1915
QY 561 VMCIIQSHPQANNEIROKLTSAVTKPPQ 589
Db 1916 AIDAATSNADVTAKQTEG--TOAINAVPQ 1942

RESULT 7
Q07290
ID Q07290 PRELIMINARY; PRT; 1822 AA.
AC Q07290;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Epf* protein.
GN EPF*.
OS Streptococcus suis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1307;
RN [1]_TaxID=1307;
RP SEQUENCE FROM N.A.
RC STRAIN=1890;
RX MEDLINE=93328288; PubMed=8335363;
RA Smith H.E., Reek F.H., Vecht U., Gielkens A.L.J., Smits M.A.;
RT "Repeats in an extracellular protein of weakly pathogenic strains of
RT Streptococcus suis type 2 are absent in pathogenic strains.";
RL Infect. Immun. 68:3318-3326(1999).
DR EMBL; X71880; CAA50714.1; -
DR PIR; S33441; S33441.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005897; Gpos YSIRK.
DR Pfam; PF04650; YSIRK signal; 1.
DR TIGRFAMs; TIGR01168; YSIRK signal; 1.
SQ SEQUENCE 1822 AA; 192632 MW; 3838960C77641D7D CRC64;

Query Match
Best Local Similarity 21.4%; Pred. No. 0.027;
Matches 166; Conservative 127; Mismatches 289; Indels 192; Gaps 34;

QY 10 IDETERTPPADUSAGLEASANKGAEQRIAGABAKPKESKTDV---ERWSILERSAVN 66
Db 1072 IDDNPNLTPE-----KESAKNAVEEAHVATA-AIDKASTPDVAQVVEEDKGV--AAIN 1122
QY 67 ALMSLADKLGIASSNSSSTSRSD-----VDSTTAT-----APTTPPTFDYKT 112
Db 1123 LITAKADAGVIAAKLADIKELEKQBAEAKAIDAATSNTEEKAIAKALQDVVDKGA 1182
QY 113 Q-----AQTAAYDTFTSTSLADIQAALV-----SLQDAVTNIXDTAATDEETAIAAE-WETK 163
Db 1183 ELEDAARVATNEIHEATTTTEKAAELAGEKSLITDGTKEARDAVELAKDELAKAIRE 1242
QY 164 NADAVKVGQITELAKYA-----SDNQ-----ILDSIGKLTSTPL 199
Db 1243 EEEATKIVEKLAEDTRKATIEDNPNLSDEDKQAEIKKLTDAVAKTLATIRDNADKRTQ--- 1299
QY 200 LQALLOSVANNKKAELKEMQDNVPVPGKTPATAQSLVDOTDATA-TQIEKQGNARD 258
Db 1300 -EAKQAALADLEKAKETQK-TADKAAIDRLTILVKDGELEATKQDAKNKIADKAAAKE 1357
QY 259 AVFAGQNASGA-----VENAKSNNSIS-----NIDSAKAAIA----- 290
Db 1358 AIASPNLTDAEKKTTTDAVDAEVAKANDAIASAATSPADVQKEEDAGVAAIAEDVLDAAK 1417
QY 291 -TAKTQIAE-----AQKKFPDPSILOAEQMWTOAEKDLKNIKPAGDSVPNPGTTVGSK 345
Db 1418 QDAKNKIADAAAAKEAIGSNPNLTDAEKKTTTDAVDAEVAKANDAIASAATSPADV--QK 1475

```

```

QY 346 QGSSIGSIRVSM---LDDAENETASTILMSGFQMIHMFNTENPDSSAAQOELAAQARA 402
Db 1476 BEDAGVAAIEDVLDAAKQDAKNKIAKESDAKSAIDANPNLTDAEKESAKKAVDADAKA 1535
QY 403 AKAGDDSSAAALADAQKALEAALGKAGQ-----QQGILNALGQIASAA----- 446
Db 1536 ATDAID--ASTSPVEAQSAEDKGVGSIADQVLDAAKQAKNIKAEVAAAKEAIDANPNL 1593
QY 447 -----VVSAGVPPAAASS-----IGSSVKQLYKTSKSTGSDYKT 480
Db 1594 SDAEKEASKKAVDADAKATTDAIDASTSPVEAQSAEDKGVGSIQDVLDAAK---QDAKN 1650
QY 481 QISAGYDAYKSINDAYGR---ARNDATRDVIN---NVSTPALTRSVPRARTEARGPE-KT 533
Db 1651 KIAKESDAKSAIDANPNLTDAEKESAKKAVDADAKAATDAIDAST--SPVEAQSAEDKG 1708
QY 534 DOALARVISGNSRTILGDVYSQVS-ALQSMVOIIGSNP-----CANNEETROKLTSAVT 585
Db 1709 VGATAKDLDAKQ--DAKNKIAKEAESAKSVISNPNLTDAEKAKEASEIDKAVEEAIV 1766
QY 586 -----KPPQFGYPYVQLSNDSTOKPIAKLESIFAEGS 617
Db 1767 LINGVRTYQELEKIKLPMAALIKDAAKVTPVD--PNLTKEKIAKIAFLKENN 1819

RESULT 8
Q9RL69
ID Q9RL69 PRELIMINARY; PRT; 2478 AA.
AC Q9RL69;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Mrp protein.
GN MRP.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]_TaxID=1280;
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=97302526; PubMed=9158773;
RA Wu S., de Lencastre H., Sali A., Tomasz A.;
RT "A phosphoglucotomase-like gene essential for the optimal expression
RT of methicillin resistance in Staphylococcus aureus: molecular cloning
RT and DNA sequencing.";
RL Microb. Drug Resist. 2:277-286(1996).
RN [2]
RC SEQUENCE FROM N.A.
RX STRAIN=COL;
RX MEDLINE=97431478; PubMed=9286983;
RA Jolly L., Wu S.W., Van Heijenoort J., de Lencastre H.,
RA Mengin-Lecreulx D., Tomasz A.;
RT "The femR315 gene from Staphylococcus aureus, the interruption of
RT which results in reduced methicillin resistance, encodes a
RT phosphoglucosamine mutase.";
RN [3]
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=99261521; PubMed=10332717;
RA Wu S., de Lencastre H.;
RT "Mrp-a new auxiliary gene essential for optimal expression of
RT methicillin resistance in Staphylococcus aureus.";
RL Microb. Drug Resist. 5:9-18(1999).
DR EMBL; Y09927; CAB55329.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos YSIRK.
DR Pfam; PF04650; YSIRK signal; 1.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; YSIRK signal; 1.
SQ SEQUENCE 2478 AA; 263031 MW; 6B9859A02D023C74 CRC64;

```

```

Query Match      7.3%; Score 231.5; DB 2; Length 2478;
Best Local Similarity 22.8%; Pred. No. 0.042;
Matches 169; Conservative 128; Mismatches 285; Indels 159; Gaps 35;

Qy 4 PIGPGPIDE-----TERTPPADLSAQGLEASANKSA-EAQRAGABAKPKESKIDS--- 54
Db 780 PLNPDTTNEEVAERINAAKVS--GVKAIEATTTAQDLERVKNIEISKIENTIDSTQT 837

Qy 55 -----VERWSILRSVAVNALMS-----LADKLG-----IASSNSSSSSTS 87
Db 838 KMDAYNEVKQAATARKAQNATVSNAATNEEVAEADAADAAKQGLHDIQVVKSKQEVADT 897

Qy 88 RSADVSTTA-TAPTPPPPTDDYKTOAQATYDT-----IFTSTSLADIQAALVSL--- 137
Db 898 KSKVLDKINAIQTCAKVKPAAD--TEVENAYNTRKQEIQNSNASTTEEKQAAVTELDTK 954

Qy 138 -QDAVTNIKDTAATD-----BETAIAAEWETKNADAVKGAQITELAKYASDNOAILDS 190
Db 955 KQEARTNL-DAANTNSDVTTAKDNSIAINQVQAATTKSDAK-AETAKASERKTAIEA 1012

Qy 191 LGKLTSPDLLQAA---LLQSVANNKAAELLKEMQDNVPVPGKT----- 231
Db 1013 MNDSTT-EEQAAKDKVDQAVVTAN--ADIDNAAANDVDNAKTNEATIAAITPDANVK 1069

Qy 232 PAIAQSLVDQDATATQIE-KDGNAIRDAYFAGQ-----NAGAVENAKSNNSI--- 279
Db 1070 PAAQAIADKVOAQETALDGNNGSTTEEKAQKQVQTEKTTADAIDAHAHTNAEVEAAK 1129

Qy 280 -----SNIDSAKAAIAT---AKTOIAEAKKFPDPSPILOEA--EQMVIQA 319
Db 1130 KAAIAKIEAIQPAITTKDNEAKAIAATKANERKTAIAQTDITABEIAAANADVNAVTOA 1189

Qy 320 EKDLNLIKPADG-SDVPNPGTIVGSKQOGSSIGSVKQYKTSKTSYDQYKTOISAGYDAYK 378
Db 1190 N---SNIEAANSQNDVQAKTGTGENSIDQVTPVVKAT-----ARNEITAILNNKLOEI 1241

Qy 379 IHMNTENPDSQAQOELAAQARAAGDDSA---ALADAQKALEAALGKAGQ 431
Db 1242 QATPDATDEEKQADAE--ANTENKANQAIISAATNAQVDEKANAEEAINAVTPKVVK 1299

Qy 432 QOGLINALGQI-ASAAVVSAGVPPAAASSIGSVKQYKTSKTSYDQYKTOISAGYDAYK 490
Db 1300 KQAAKDEIDQATQTNVINNDQNAATTEKEAAIQQL-----ATAVTDKANNITATD--- 1352

Qy 491 SINDAYGRANDATRDVINNVSTPALTRSPRARTPEKTPDQALRVISGNSRTIGD 550
Db 1353 --DNGVQAK-DAGKNSIQSTQ-----PATAVKSNKNDVDQAVTTQONQIDNTTGA 1401

Qy 551 VYSQVSAQSVQMIIQSNPOANNEIROKLTSAVTKPPQFGYPYVQ-LSNDSTOKFIKL 609
Db 1402 TTEKNAAKDL--VLKAKAYQDILNAQTINDVTQIKDOAVADIQGITADTTIKDVAK- 1458

Qy 610 ESLFAEGSRTAAEIKALSFT 630
Db 1459 ----DELATKANEQKALIAQT 1475

RESULT 9
Q9LCH2
ID Q9LCH2 PRELIMINARY; PRT; 2478 AA.
AC Q9LCH2;
DT 01-OCT-2000 (TreeBrel. 15, Created)
DT 01-OCT-2000 (TreeBrel. 15, Last sequence update)
DT 01-JUN-2003 (TreeBrel. 24, Last annotation update)
DE FmtB.
GN FMTB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=20348625; PubMed=10896508;

```

```

RA Komatsuzawa H., Ohta K., Sugai M., Fujiwara T., Glanzmann P.,
RA Berger-Bachi B., Suganaka H.;
RT "Trf51-mediated insertional inactivation of the fntB gene encoding a
RT cell wall-associated protein abolishes methicillin resistance in
RT Staphylococcus aureus.";
RL J. Antimicrob. Chemother. 45:421-431 (2000).
DR EMBL; AB025716; BAA93438.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR005877; Gpos_Y5IRK.
DR Pfam; PF04650; Y5IRK signal; 1.
DR TIGRFAMs; TIGR01167; LPX7C anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK signal; 1.
SQ SEQUENCE 2478 AA; 262995 MW; 1C118BEBE0DB03B34 CRC64;

Query Match      7.2%; Score 229.5; DB 2; Length 2478;
Best Local Similarity 22.8%; Pred. No. 0.052;
Matches 169; Conservative 127; Mismatches 286; Indels 159; Gaps 35;

Qy 4 PIGPGPIDE-----TERTPPADLSAQGLEASANKSA-EAQRAGABAKPKESKIDS--- 54
Db 780 PLNPDTTNEEVAERINAAKVS--GVKAIEATTTAQDLERVKNIEISKIENTIDSTQT 837

Qy 55 -----VERWSILRSVAVNALMS-----LADKLG-----IASSNSSSSSTS 87
Db 838 KMDAYNEVKQAATARKAQNATVSNAATNEEVAEADAADAAKQGLHDIQVVKSKQEVADT 897

Qy 88 RSADVSTTA-TAPTPPPPTDDYKTOAQATYDT-----IFTSTSLADIQAALVSL--- 137
Db 898 KSKVLDKINAIQTCAKVKPAAD--TEVENAYNTRKQEIQNSNASTTEEKQAAVTELDTK 954

Qy 138 -QDAVTNIKDTAATD-----BETAIAAEWETKNADAVKGAQITELAKYASDNOAILDS 190
Db 955 KQEARTNL-DAANTNSDVTTAKDNSIAINQVQAATTKSDAK-AETAKASERKTAIEA 1012

Qy 191 LGKLTSPDLLQAA---LLQSVANNKAAELLKEMQDNVPVPGKT----- 231
Db 1013 MNDSTT-EEQAAKDKVDQAVVTAN--ADIDNAAANDVDNAKTNEATIAAITPDANVK 1069

Qy 232 PAIAQSLVDQDATATQIE-KDGNAIRDAYFAGQ-----NAGAVENAKSNNSI--- 279
Db 1070 PAAQAIADKVOAQETALDGNNGSTTEEKAQKQVQTEKTTADAIDAHAHTNAEVEAAK 1129

Qy 280 -----SNIDSAKAAIAT---AKTOIAEAKKFPDPSPILOEA--EQMVIQA 319
Db 1130 KAAIAKIEAIQPAITTKDNEAKAIAATKANERKTAIAQTDITABEIAAANADVNAVTOA 1189

Qy 320 EKDLNLIKPADG-SDVPNPGTIVGSKQOGSSIGSVKQYKTSKTSYDQYKTOISAGYDAYK 378
Db 1190 N---SNIEAANSQNDVQAKTGTGENSIDQVTPVVKAT-----ARNEITAILNNKLOEI 1241

Qy 379 IHMNTENPDSQAQOELAAQARAAGDDSA---ALADAQKALEAALGKAGQ 431
Db 1242 QATPDATDEEKQADAE--ANTENKANQAIISAATNAQVDEKANAEEAINAVTPKVVK 1299

Qy 432 QOGLINALGQI-ASAAVVSAGVPPAAASSIGSVKQYKTSKTSYDQYKTOISAGYDAYK 490
Db 1300 KQAAKDEIDQATQTNVINNDQNAATTEKEAAIQQL-----ATAVTDKANNITATD--- 1352

Qy 491 SINDAYGRANDATRDVINNVSTPALTRSPRARTPEKTPDQALRVISGNSRTIGD 550
Db 1353 --DNGVQAK-DAGKNSIQSTQ-----PATAVKSNKNDVDQAVTTQONQIDNTTGA 1401

Qy 551 VYSQVSAQSVQMIIQSNPOANNEIROKLTSAVTKPPQFGYPYVQ-LSNDSTOKFIKL 609
Db 1402 TTEKNAAKDL--VLKAKAYQDILNAQTINDVTQIKDOAVADIQGITADTTIKDVAK- 1458

Qy 610 ESLFAEGSRTAAEIKALSFT 630
Db 1459 ----DELATKANEQKALIAQT 1475

RESULT 10

```

```

O99QR6
ID Q99QR6 PRELIMINARY; PRT; 2481 AA.
AC Q99QR6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FmtB protein.
GN FMTB(MRP) OR FMTB OR SAV2160 OR SAI1964.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacilliales; Staphylococcus.
OX NCBI_TaxID=158978, 158979;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U I., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanesaka M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB58322.1; -.
DR EMBL; AP003364; BAB43253.1; -.
DR PIR; D90011; D90011.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016774; F:phosphotransferase activity, carboxyl group. .; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0016310; P:phosphorylation; IEA.
DR InterPro; IPR000890; Acetate kin.
DR InterPro; IPR005877; Gpos YSIRK.
DR InterPro; IPR001899; Gram_pos anchor.
DR Pfam; PF04650; YSIRK_signal; 1.
DR TIGRfam; TIGR01167; LPXTG_anchor; 1.
DR TIGRfam; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS01075; ACETATE_KINASE_1; 1.
KW Complete proteome.
SQ SEQUENCE 2481 AA; 263765 MW; ELEAB99B81665E8 CRC64;

Query Match 7.0%; Score 223.5; DB 16; Length 2481;
Best Local Similarity 22.1%; Pred. No. 0.097;
Matches 163; Conservative 129; Mismatches 293; Indels 153; Gaps 31;

QY 4 PIGPGPIE-----TERTPPADLSAQGLEASAAKSA-EAQRAGAAKPKESKTDSDS--- 54
DB 763 PLNPFTTNEEVAEATERINAQVSV--GVKAIEATTAQDLERVKNEIFKINENITDSTQT 820
QY 55 -----VERWSILRSVAVNALMSLADKLGIASSNSSSSSTSRSD-----VDSTTTATAPT 101
DB 821 KMDAYKEVRQAATARKAQAATVSNATDEEVAEANAADAAQTEGLHDIIQVVKSQQEVADT 880
QY 102 PP-----PPTFDYKTOQATVDT-----IPTSTSLADIQAULVSL--- 137
DB 881 KAKVLDKINAIGTOAKVKAAD---TEVENAYNTRKQEIQNSNASTTEKEAAYTDLAK 937
QY 138 -QDAVTNKDITAATD-----EETAIAAEWEKNDADAVKGAQITELAKYASDNQAILDS 190
DB 938 QKEATNL-DAANTNSDVTAKDNGIAINQVAATTKSKDAK-AEIAQKASERTALIEA 995
QY 191 LKGLT-----SFDLQALLOSVAANKA-----AELKEMQDNFVPGKT--PAI 234
DB 996 MNDSTTEEQAAKQVDQAVVTANADIDNATANTDNDNAKTNTNEATIAAITPDANVKFAA 1055
QY 235 AQLVDQDTATATQIE-KDGNRAIRDAYFAGQN-----ASGAVNAKNSNI----- 279
DB 1056 KOAIAKRVQAQOETAIDANNSTTEKEAKQVQTEKTAADAADAIDAAHSNVEVEAKNAE 1115

```



```
Db 1737 ESSAATEIYQETQIASSIAAGALGEAHTSELAEASSSSSAASAAAAEQSLYDTSSA 1796
QY 555 VQALQSV-----MGIQSNPANNNEEROKLTSAVTTPPOFGYPYVOLSNDSTQK 604
Db 1797 ASASASSDFIASDITNQSLSVNSASSAAEESVSQVDETYQ---NFDQYSSISASA 1853
QY 605 FIAKLESFAEGSRVTAETK---ALSFETNSLFIQQVLVNIIGSLYS 647
Db 1854 SAAQSSEIYQVSSSSAAASTSSAASSLETSGTVAESGSGTAASSSYA 1900

RESULT 15
Q86RN1 PRELIMINARY; PRT; 842 AA.
ID Q86RN1;
AC Q86RN1;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Merozoite surface protein 3 alpha (Fragment).
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22346636; PubMed=12458823;
RA Rayner J.C., Corredor V., Feldman D., Ingravallo P., Iderabdullah F.,
RA Galinski M.R., Barnwell J.W.;
RT "Extensive polymorphism in the plasmodium vivax merozoite surface coat
RT protein MSP-3alpha is limited to specific domains."
RL Parasitology 125:393-405(2002).
DR EMBL; AF491951; AAC20882.1; -.
FT NON TER 842
SQ SEQUENCE 842 AA; 89463 MW; E0B13FA7D9C65169 CRC64;
```

```
Query March 6.7%; Score 214; DB 5; Length 842;
Best Local Similarity 19.9%; Pred. No. 0.066;
Matches 137; Conservative 107; Mismatches 291; Indels 154; Gaps 22;

QY 27 EASAANKSAEAPQRIAG-----AEAKPKESKTDVSVERWSILRSAY--NA 67
Db 163 QAAIAEKADAAEAGAKENKLDVKSQVIAEEASTKAKDKTEAEIAVEIVKAVAKEE 222
QY 68 LMSLADKLGTASNSSSSTSRSDVSTTATATPPPPFDYKTAQATRYDTIFTSTSL 127
Db 223 AQKASDEAQACEKAKAHAKAQKASDTTKTVET-----FKTNAEAAKNAKEKAGN 274
QY 128 ADIQAALVSLQDAVTNFKDTAATDETA-----IAEWETKNADAVKVGAOITELAKY 180
Db 275 ANKATEAESANELSVAKQAKADAEAAKEAKEQVKAETAAEVAKAKVAKAEAEAEQK 334
QY 181 ASDNQAILDSGLKLTSPDLLQAALLOSVAANNKAAELLKEMQDNVVPVFGKTPAIAQSLVD 240
Db 335 ABEAKKIVDKIAQDTKVPEQAQAAEFATETVKATTAATEAGKNAQAEKLPAAEET--- 391
QY 241 QTATATQIEKGNATRDAYFAGNAGSAGVENAKSNNSISNIDSAKAAIATAKTQIAEAQ 300
Db 392 -----SDAVKGKADAEAKAAGEAKKASITETI-----AIEVAKAEVLNAE 431
QY 301 KKPDPSPILQEAQOMVIQAEKOLKNIKPA-----DGSVDVPNGTGVGSKQGGSSIG 352
Db 432 VK----KTAQAEAKDAEAEQAEKAAAEAEAKTHGEKAEKVESTKAHSDKAQENKK 487
QY 353 SIRVSMILLDDAENETASILMSGFQMIHMFNTENPDSQA----- 391
Db 488 AKEAS---DEAENRAVDALAEAYAVEAHLARTKNAEBSAKSATDLSKLEAKEEAIDAN 544
QY 392 -AQCELAQAAPAAKAGDSDSAAALA--DAQALEALGKAGQOQGIINALGQIASAAVV 448
Db 545 IAHQKWLKATQAATTAKETKEAAKVAAEKTKETATAAKLKAAKAEA-KKAETEAVKAAVE 603
QY 449 SAGVPPAA---ASSIGSVKLYKT-----SKSTGSDYKTIQISAGVDAYKINSIDAYGR 498
```

```
Db 604 ARAAAEAKQEAQKVGAS-KEPQETKNKANYEAETGNEAK-----KAEDAAEEVKEAAKK 658
QY 499 ARNDATRDVINNVSTPALTRSPVPRTAREPKEKTDQALARVIGNSRTLGDVYSQVSAL 558
Db 659 A-NEATD---ANVARSEADKAIAAAKAKAKAREKAAAYGLLKTKN----- 698
QY 559 QSVMOIIOQSNPQ-----ANNEEIRQKLTSAVTKPPQFGYPYVOLSNDSTQKFIAKLES 612
Db 699 QYVLEPLDISPSADNITSKEEQVKEVED-----QDEDNSE---AEVEEA 742
QY 613 FAEGSRTAAEIKALSFTETNSLFIQQVLVN 641
Db 743 LPNGSGAQEEDUNLEMDDEEVEENVATN 771
```

Search completed: March 24, 2004, 05:59:09
Job time : 61.0374 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:04 ; Search time 713.142 Seconds
(without alignments)
11032.393 Million cell updates/sec

Title: US-10-608-559-3

Perfect score: 1852
Sequence: 1 atgagtgcagataagct.....gatcttcataataaaga 1852

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002s.*

7: geneseqn2003as.*

8: geneseqn2003bs.*

9: geneseqn2003cs.*

10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1852	100.0	1852	3 AAD02064	Aad02064 5'-trunca
2	1848.8	99.8	2156	3 AAD02063	Aad02063 Chlamydia
3	1836.2	99.1	110000	2 AAX91990_08	Continuation (9 of
4	1748.8	94.4	1956	4 AAS57031	Aas57031 C. pneumo
5	1748.8	94.4	1956	6 ABL92612	Ab192612 Chlamydia
6	1748.8	94.4	1956	6 ABL91190	Ab191190 Chlamydia
7	1748.8	94.4	1956	9 ADD42825	Add42825 Chlamydia
8	1161.8	62.7	2238	3 AAD02066	Aad02066 C. pneumo
9	1151	62.1	1456	3 AAD02065	Aad02065 3'-trunca
10	187.6	10.1	110000	2 AAZ01425_07	Continuation (8 of
11	187.6	10.1	110000	2 AAZ01425_08	Continuation (9 of
12	187.2	10.1	1537	4 AAS63292	Aas63292 C. tracho
13	187.2	10.1	1537	4 AAS6196	Aas6196 Chlamydia
14	187.2	10.1	1537	6 ABL92425	Ab192425 Chlamydia
15	186	10.0	1944	9 ADD43867	Add43867 Chlamydia
16	185.4	10.0	1941	6 ABL92619	Ab192619 Chlamydia
17	184	9.9	1171	4 AAS56981	Aas56981 C. tracho
18	184	9.9	1171	9 ADD42775	Add42775 Chlamydia
19	184	9.9	1834	4 AAS57006	Aas57006 C. tracho
20	184	9.9	1834	9 ADD42800	Add42800 Chlamydia
21	183	9.9	1983	4 AAS56996	Aas56996 C. tracho
22	183	9.9	1983	9 ADD42790	Add42790 Chlamydia
23	148.4	8.0	150	2 AAV16207	Aav16207 Part of t

24	51.6	2.8	2205	8	ADA28999	Ada28999 DNA encod
25	49.8	2.7	2000	7	ADA71938	Ada71938 Rice gene
26	49	2.6	896	2	AAQ68903	PspA prot
27	49	2.6	946	2	AAQ33124	Streptoco
28	49	2.6	957	2	AAT61726	Streptoco
29	49	2.6	1860	7	ACA49648	Prokaryot
30	49	2.6	1990	2	AAT61725	Streptoco
31	49	2.6	1990	7	ABX95373	S. pneumo
32	49	2.6	2085	2	AAQ78131	Pneumococ
33	49	2.6	2085	2	AAT08979	DNA encod
34	49	2.6	2085	2	AAT07178	Pneumococ
35	49	2.6	2085	2	AAV33264	Pneumococ
36	49	2.6	2085	2	AAV33470	Streptoco
37	49	2.6	2085	2	AAZ25063	Streptoco
38	49	2.6	2085	2	AAV84069	DNA encod
39	49	2.6	2085	2	AAV84069	DNA encod
40	47.4	2.6	956	7	ABX95374	S. pneumo
41	47.4	2.6	2085	2	AAQ28674	Abx28674 PspA gene
42	47.2	2.5	110000	6	ABA90521_07	Continuation (8 of
43	47	2.5	9373	6	ABL33291	Human imm
44	47	2.5	9373	6	ABK31317	Signal tr
45	47	2.5	9373	6	ABL70560	Chemicall

ALIGNMENTS

RESULT 1
AAD02064
ID AAD02064 standard; DNA; 1852 BP.
AC AAD02064;
XX AAD02064;
DT 26-MAR-2001 (first entry)
XX 5'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.
XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; truncation mutant; ds.
XX Chlamydophila pneumoniae.
OS Synthetic.
FH Key Location/Qualifiers
FT CDS 1..1752
FT /tag= a
FT /product= "5'-truncated Chlamydia pneumoniae 76KDa
FT protein"
FT /transl_except= (pos:1489..1491, aa:lle)

WO200066739-A2.
09-NOV-2000.
03-MAY-2000; 2000WO-CA0000511.
03-MAY-1999; 99US-0132270P.
30-JUN-1999; 99US-0141276P.
(AVET) AVENTIS PASTEUR LTD.
Murdin AD, Oomen RP, Wang J, Dunn P;
WPI; 2000-687542/67.
P-PSDB; AAY71955.
Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
for vaccinating against Chlamydia infections.
Claim 2b; Page 97-99; 112pp; English.

CC The present sequence is a DNA coding for 5'-truncated Chlamydia
CC pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the
CC diagnosis, prevention and treatment of C. pneumoniae infections (e.g.
CC pneumonia, upper respiratory tract disease, bronchitis, sinusitis and
CC acute respiratory disease such as cough, sore throat, hoarseness, fever;
CC and abnormal chest sounds on auscultation). C. pneumoniae sequence is
CC also used as vaccines for immunising humans against diseases caused by C.
CC pneumoniae
XX

SQ Sequence 1852 BP; 578 A; 409 C; 407 G; 458 T; 0 U; 0 Other;

Query Match 100.0%; Score 1852; DB 3; Length 1852;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGTCTGGCAGATTAAGCTGGTATTGCTTCTAGTAACAGCTGCTTCTACTAGCAGA	60
DB	1	ATGAGTCTGGCAGATTAAGCTGGTATTGCTTCTAGTAACAGCTGCTTCTACTAGCAGA	60
QY	61	TCTGAGACGTGGACTCAACGACAGCGACCGACCTACGCTTCTCCACCGCTTTGAT	120
DB	61	TCTGAGACGTGGACTCAACGACAGCGACCGACCTACGCTTCTCCACCGCTTTGAT	120
QY	121	GATTATAGACTCAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT	180
DB	121	GATTATAGACTCAGCGCAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT	180
QY	181	GACATACAGGCTGCTTTGGTGGCTCCAGAGTGTGTCACTAATATAAGATACACGG	240
DB	181	GACATACAGGCTGCTTTGGTGGCTCCAGAGTGTGTCACTAATATAAGATACACGG	240
QY	241	GCTACTGATAGGAACCGCAATCGCTCGGAGTGGGAACTAAGAAATGCCGATGCA	300
DB	241	GCTACTGATAGGAACCGCAATCGCTCGGAGTGGGAACTAAGAAATGCCGATGCA	300
QY	301	AAAGTTGGCGCGCAATACAGAAATAGCGAAATAGCTTCGGATAACCAAGCGATCTT	360
DB	301	AAAGTTGGCGCGCAATACAGAAATAGCGAAATAGCTTCGGATAACCAAGCGATCTT	360
QY	361	GACTCTTTAGGTAACTGACATCTCTTTCGACCTCTTACAGGCTGCTTCTTCTCAATCTGTA	420
DB	361	GACTCTTTAGGTAACTGACATCTCTTTCGACCTCTTACAGGCTGCTTCTTCTCAATCTGTA	420
QY	421	GCAACAAATACAAAGCGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTATGCCA	480
DB	421	GCAACAAATACAAAGCGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTATGCCA	480
QY	481	GGGAAACGCTGCAATCTTCTTGTAGTATCAGACAGATGCTTACAGCGACACAG	540
DB	481	GGGAAACGCTGCAATCTTCTTGTAGTATCAGACAGATGCTTACAGCGACACAG	540
QY	541	ATAGAGAAGATGGAATTCGATAGGATGATATTTTGCAGGACAGAACGCTAGTGGG	600
DB	541	ATAGAGAAGATGGAATTCGATAGGATGATATTTTGCAGGACAGAACGCTAGTGGG	600
QY	601	GCTGTAGAAATGCTAAATCTTAATACAGTATACAGTATACAGTATACAGTAAAGCAGCA	660
DB	601	GCTGTAGAAATGCTAAATCTTAATACAGTATACAGTATACAGTATACAGTAAAGCAGCA	660
QY	661	ATCGCTACTGCTTAAGACCAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCCAATT	720
DB	661	ATCGCTACTGCTTAAGACCAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTCCAATT	720
QY	721	CTTCAAGACGGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTTAAATAATCAAACT	780
DB	721	CTTCAAGACGGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTTAAATAATCAAACT	780
QY	781	GCAGATGGTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACAGGA	840
DB	781	GCAGATGGTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAGCAACAGGA	840
QY	841	AGTACTATTGGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAATGAGACGCT	900
DB	841	AGTACTATTGGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAATGAGACGCT	900

RESULT 2

AAD02063

ID AAD02063 standard; DNA; 2156 BP.

XX

AC

XX

DT	15-SEP-2003	(revised)
DT	26-MAR-2001	(first entry)
XX		
XX	Chlamydia pneumoniae	76 kDa full-length protein encoding DNA.
XX		
XX	76 kDa protein; bactericidal; diagnosis; prevention; treatment;	
KW	pneumonia; upper respiratory tract disease; bronchitis; sinusitis;	
KW	acute respiratory disease; cough; sore throat; hoarseness; fever;	
KW	vaccine; immunisation; ds.	
XX		
XX	Chlamydophila pneumoniae.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	101..2056
FT	/tag= a	
FT	/product= "Chlamydia pneumoniae 76kDa protein"	
XX	WO2000066739-A2.	
PN		
PD	09-NOV-2000.	
XX		
XX	03-MAY-2000; 2000WO-CA000511.	
XX		
PR	03-MAY-1999; 99US-0132270P.	
PR	30-JUN-1999; 99US-0141276P.	
XX		
PA	(AVET) AVENTIS PASTEUR LTD.	
PI	Murdin AD, Oomen RP, Wang J, Dunn P;	
XX		
DR	WPI; 2000-687542/67.	
DR	P-PSDB; AAY1954.	
XX		
PT	Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful	
PT	for vaccinating against Chlamydia infections.	
XX		
PS	Claim 2a; Fig 1; 112pp; English.	
XX		
CC	The present sequence is a DNA coding for Chlamydia pneumoniae 76 kDa	
CC	protein. C. pneumoniae 76 kDa protein is used in the diagnosis,	
CC	prevention and treatment of C. pneumoniae infections (e.g. pneumonia,	
CC	upper respiratory tract disease, bronchitis, sinusitis and acute	
CC	respiratory disease such as cough, sore throat, hoarseness, fever; and	
CC	abnormal chest sounds on auscultation). C. pneumoniae sequence is also	
CC	used as vaccines for immunising humans against diseases caused by C.	
CC	pneumoniae. (Updated on 15-SEP-2003 to standardise OS field)	
XX		
SQ	Sequence 2156 BP; 672 A; 451 C; 471 G; 552 T; 0 U; 0 Other;	
	Query Match	99.8%; Score 1848.8; DB 3; Length 2156;
	Best Local Similarity	99.9%; Pred. No. 0;
	Matches 1850; Conservative	0; Mismatches 2; Indels 0; Gaps 0
QY	1 ATGAGTCTGGCAGATAAGCTGGGTATTCCTTTCTAGTAACAGCTCGTCTTCTACTAGCAGA	60
Db	305 ATGAGTCTGGCAGATAAGCTGGGTATTCCTTTCTAGTAACAGCTCGTCTTCTACTAGCAGA	364
QY	61 TCTGCAGACGTGGACTCAACACACAGCACGCCGACCCTAGCCCTCTCCACCCACGTTTGAT	120
Db	365 TCTGCAGACGTGGACTCAACACACAGCACGCCGACCCTAGCCCTCTCCACCCACGTTTGAT	424
QY	121 GATTATTAAGACTCAAGCGCAACACAGCTTAGCATCTACTTTTACTCTCAACATCACTAGCT	180
Db	425 GATTATTAAGACTCAAGCGCAACACAGCTTAGCATCTACTTTTACTCTCAACATCACTAGCT	484
QY	181 GACATACAGGCTGCTTTTGGTGGAGCCTCCAGGATGCTGCTAATAATATAAGGATACAGCG	240
Db	485 GACATACAGGCTGCTTTTGGTGGAGCCTCCAGGATGCTGCTAATAATATAAGGATACAGCG	544
QY	241 GCTACTGATGAGGAACCGGAATGCTCGGAGTGGGAAACTAAGAANTCCGATCGAGTT	300
Db	545 GCTACTGATGAGGAACCGGAATGCTCGGAGTGGGAAACTAAGAANTCCGATCGAGTT	604

```
Db 1685 GGACGAGAAACAGATCAAGCCCTCGCTAGGGTGAATTTCTGGCAATAGCAGAACTCTT 1744
Qy 1441 GGAGATGCTATAGTCAAGTTTCGGCACTACAACTCTGTAATGCAGATCACTCAGTCGAAT 1500
Db 1745 GGAGATGCTATAGTCAAGTTTCGGCACTACAACTCTGTAATGCAGATCACTCAGTCGAAT 1804
Qy 1501 CCTCAGGGAATATAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCCCTCCA 1560
Db 1805 CCTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCCCTCCA 1864
Qy 1561 CAGTTTGGCTATCTTATGTGCAACTTCTTAATGACTCTACACAGAAGTTTCATAGCTAAA 1620
Db 1865 CAGTTTGGCTATCTTATGTGCAACTTCTTAATGACTCTACACAGAAGTTTCATAGCTAAA 1924
Qy 1621 TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAGACACCTTCCTTT 1680
Db 1925 TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAGACACCTTCCTTT 1984
Qy 1681 GAAACGAACTCTTGTATTCAGAGGTGCTGTGTAATATCGGCTCTCTATATCTCTGGT 1740
Db 1985 GAAACGAACTCTTGTATTCAGAGGTGCTGTGTAATATCGGCTCTCTATATCTCTGGT 2044
Qy 1741 TATCTCCAATAACACACCTAAAGTTCGTTGGAGAGATTATATGCTGCTTTGGTAAGG 1800
Db 2045 TATCTCCAATAACACACCTAAAGTTCGTTGGAGAGATTATATGCTGCTTTGGTAAGG 2104
Qy 1801 CCTTTGTTGGGCTTACCAACACACTAGAACGATTCCTCAATAAATAAAGA 1852
Db 2105 CCTTTGTTGGGCTTACCAACACACTAGAACGATTCCTCAATAAATAAAGA 2156

RESULT 3
AAX91990_08/c
Continuation (9 of 13) of AAX91990 from base 800001 (Nucleotide sequence of the complete
WP Sequence split into 13 fragments LOCUS AAX91990 Accession Aax91990
WP Fragment Name Begin End
WP AAX91990_00 1 110000
WP AAX91990_01 100001 210000
WP AAX91990_02 200001 310000
WP AAX91990_03 300001 410000
WP AAX91990_04 400001 510000
WP AAX91990_05 500001 610000
WP AAX91990_06 600001 710000
WP AAX91990_07 700001 810000
WP AAX91990_08 800001 910000
WP AAX91990_09 900001 1010000
WP AAX91990_10 1000001 1110000
WP AAX91990_11 1100001 1210000
WP AAX91990_12 1200001 1230025

Query Match 99.1%; Score 1836.2; DB 2; Length 110000;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1849; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTCTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
Db 28394 ATGAGTCTGGCAGATAAGCTGGGTATGCTCTCTAGTAACAGCTCGTCTTCTACTAGCAGA 28335
Qy 61 TCTCAGACGCTGAGCTCAACGACAGCGACCGACCTACGCTCTCTCCACCCACGTTTGAT 120
Db 28334 TCTCAGACGCTGAGCTCAACGACAGCGACCGACCTACGCTCTCTCCACCCAGCTTTGAT 28275
Qy 121 GATTATAGACTCAAGCGCAACAGCTTACATATCTTCTTACTCAATCACTAGCT 180
Db 28274 GATTATAGACTCAAGCGCAACAGCTTACATATCTTCTTACTCAATCACTAGCT 28215
Qy 181 GACATACAGGCTGCTTTGGTGAGCTCTCAGAGTCTGCTCAATAATAAAGGATACAGCG 240
Db 28214 GACATACAGGCTGCTTTGGTGAGCTCTCAGAGTCTGCTCAATAATAAAGGATACAGCG 28155
Qy 241 GCTTACTGATAGGAAACCGCAATCGCTCGGAGTGGGAAACATAAGATGCGGATGCGAGTT 300
```

```
Db 28154 GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGATGCCGATGCAGTT 28095
Qy 301 AAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGGATCTTT 360
Db 28094 AAAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGGATCTTT 28035
Qy 361 GACTCTTTAGGTAACATGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTTCCAAATCTGTA 420
Db 28034 GACTCTTTAGGTAACATGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTTCCAAATCTGTA 27975
Qy 421 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db 27974 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 27915
Qy 481 GGGAAAACGCTCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
Db 27914 GGGAAAACGCTCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 27855
Qy 541 ATAGAGAAAGATGGAATGGGATTAGGATGCATATTTTGCAGGACAGAACGCTAGTGA 600
Db 27854 ATAGAGAAAGATGGAATGGGATTAGGATGCATATTTTGCAGGACAGAACGCTAGTGA 27795
Qy 601 GCTGTAGAAAATGCTTAAATCTAATAACAGTATTAAGCAACATAGATTTCAGCTAAAGCAGCA 660
Db 27794 GCTGTAGAAAATGCTTAAATCTAATAACAGTATTAAGCAACATAGATTTCAGCTAAAGCAGCA 27735
Qy 661 ATCGCTACTGCTAAGACACAATAATGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAATT 720
Db 27734 ATCGCTACTGCTAAGACACAATAATGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAATT 27675
Qy 721 CTTTCAAGAGCGGAAACAATGTTTACAGGCTGAGAAAGATCTTAAATAATATCAAACT 780
Db 27674 CTTTCAAGAGCGGAAACAATGTTTACAGGCTGAGAAAGATCTTAAATAATATCAAACT 27615
Qy 781 GCAGATGGTTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAGGA 840
Db 27614 GCAGATGGTTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAGGA 27555
Qy 841 AGTAGTATTTGGTAGTATTCGTGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 900
Db 27554 AGTAGTATTTGGTAGTATTCGTGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 27495
Qy 901 TCCATTTTGTATGCTGGGTTTTCGTAGATGATTCACATGTTCAATACGGAATAATCTGAT 960
Db 27494 TCCATTTTGTATGCTGGGTTTTCGTAGATGATTCACATGTTCAATACGGAATAATCTGAT 27435
Qy 961 TCTCAAGCTGCCCAACAGAGAGCTCGCAGCACAAAGCTAGACGCGAAAGCCGCTGGAGAT 1020
Db 27434 TCTCAAGCTGCCCAACAGAGAGCTCGCAGCACAAAGCTAGACGCGAAAGCCGCTGGAGAT 27375
Qy 1021 GACAGTGTCTGCTCAGGCTGGCAGATGCTCAGAAAGCTTTAGAACGGGCTCTAGGTAA 1080
Db 27374 GACAGTGTCTGCTCAGGCTGGCAGATGCTCAGAAAGCTTTAGAACGGGCTCTAGGTAA 27315
Qy 1081 GCTGGGGCAACAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTG-TTGT 1139
Db 27314 GCTGGGGCAACAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGT 27255
Qy 1140 GAGCGGAGGATTTCTTCCGCTGSCAGCAAGTTTCTATAGGGTCACTCTGTAAAAACAGCTTTA 1199
Db 27254 GAGCGGAGGATTTCTTCCGCTGSCAGCAAGTTTCTATAGGGTCACTCTGTAAAAACAGCTTTA 27195
Qy 1200 CAAGAGCTCAAAATCTACAGGTTCTGATTTATTAACACAGATATCAGCAGGTTATGATGC 1259
Db 27194 CAAGAGCTCAAAATCTACAGGTTCTGATTTATTAACACAGATATCAGCAGGTTATGATGC 27135
Qy 1260 TTACAAATCCATCAATGATGCCCTATGGTAGGGCAGCAAAATGATGCCACTCGTGTATGAT 1319
Db 27134 TTACAAATCCATCAATGATGCCCTATGGTAGGGCAGCAAAATGATGCCACTCGTGTATGAT 27075
Qy 1320 AAACAATGTAAAGTACCCTCGCTCTCAAGATCCCGTTCTTAGAGCAGCAACAGAGCTCG 1379
Db 27074 AAACAATGTAAAGTACCCTCGCTCTCAAGATCCCGTTCTTAGAGCAGCAACAGAGCTCG 27015
```

QY 1380 AGGACAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCT 1439
 DB |||||
 DB 27014 AGGACAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCT 26955
 QY 1440 TGGAGATGCTATAGTCAAGTTTCGGCACTACAACTGTGAATGCAATCACTCAGTCGAA 1499
 DB |||||
 DB 26954 TGGAGATGCTATAGTCAAGTTTCGGCACTACAACTGTGAATGCAATCACTCAGTCGAA 26955
 QY 1500 TCTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCTCC 1559
 DB |||||
 DB 26894 TCTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCTCC 26835
 QY 1560 ACAGTTGGCTATCTTATGTGCACTTCTAATGACTCTACACAGAGTTGATAGCTAA 1619
 DB |||||
 DB 26834 ACAGTTGGCTATCTTATGTGCACTTCTAATGACTCTACACAGAGTTGATAGCTAA 26775
 QY 1620 ATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTTCCTT 1679
 DB |||||
 DB 26774 ATTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTTCCTT 26715
 QY 1680 TGAACGAACTCCTTGTATTATCAGAGGTGCTGTCATATCGGCTCTCTATATTCGG 1739
 DB |||||
 DB 26714 TGAACGAACTCCTTGTATTATCAGAGGTGCTGTCATATCGGCTCTCTATATTCGG 26655
 QY 1740 TTATCTCCAATAACACACCTAAGTGTTCGTTTGGAGAGATTTATGTGCTTGGTAA 1799
 DB |||||
 DB 26654 TTATCTCCAATAACACACCTAAGTGTTCGTTTGGAGAGATTTATGTGCTTGGTAA 26595
 QY 1800 GCCTTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 1852
 DB |||||
 DB 26594 GCCTTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATAATAAAGA 26542

RESULT 4

AA557031
 ID AA557031 standard; DNA; 1956 BP.

AC AA557031;

DT 11-SEP-2003 (revised)

DT 16-JAN-2002 (first entry)

DE C. pneumoniae DNA encoding the CT622 homologue CPr0728.

XX Chlamydia; ds; sexually transmitted disease; PID; antibacterial;

KW pelvic inflammatory disease; antigen; trachoma; gynecological;

KW acute respiratory tract infection; atherosclerosis; male infertility;

XX coronary heart disease

OS Chlamydophila pneumoniae.

XX WO2000181379-A2.

EN 01-NOV-2001.

PD 23-APR-2001; 2001WO-US013081.

XX 21-APR-2000; 2000US-0198853P.

PR 20-JUL-2000; 2000US-021952P.

XX (CORI), CORIXA CORP.

PA Bhatia A, Probst-P, Stromberg EJ;

XX WPI; 2001-616771/71.

DR P-PSDB; AAU38899.

XX New polynucleotide for treating Chlamydia infections encodes a

PT polynucleotides containing an immunogenic portion of a Chlamydia antigen.

XX PS Disclosure; Page 143; 208pp; English.

XX

CC The invention relates to isolated polynucleotide encoding at least a
 CC partial Chlamydia protein which is an antigenic fragment, or the
 CC complements, fragments, homologues and variants, and antibodies raised
 CC against the antigenic proteins (or fragments). The nucleic acids,
 CC proteins and antibodies are used to diagnose and treat Chlamydia
 CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
 CC disease (PID), acute respiratory tract infection, trachoma,
 CC atherosclerosis and coronary heart disease) in a patient, and in the
 CC treatment of male infertility. The compounds of the invention are also
 CC useful for detecting the presence of Chlamydia in a patient, and
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC The present sequence encodes a Chlamydia antigen. (Updated on 11-SEP-2003
 CC to standardise OS field)

XX
 SQ Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

Query Match 94.4%; Score 1748.8; DB 4; Length 1956;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
 DB |||||
 DB 205 ATGAGTCTGGCAGATAAGCTGGGTATTCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264
 QY 61 TCTGCAGAGCTGGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCAGCTTTCAT 120
 DB |||||
 DB 265 TCTGCAGAGCTGGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCAGCTTTCAT 324
 QY 121 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTCTCTCAACATCATAGCT 180
 DB |||||
 DB 325 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTCTCTCAACATCATAGCT 384
 QY 181 GACATACAGGCTGCTTTGGTGAGCCTCCAGATGCTGCTCACTAATAATAAGGATACACG 240
 DB |||||
 DB 385 GACATACAGGCTGCTTTGGTGAGCCTCCAGATGCTGCTCACTAATAATAAGGATACACG 444
 QY 241 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAACTAAGATCGCGATGCGATT 300
 DB |||||
 DB 445 GCTACTGATGAGGAACCGCAATCGCTCGGAGTGGGAACTAAGATCGCGATGCGATT 504
 QY 301 AAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT 360
 DB |||||
 DB 505 AAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT 564
 QY 361 GACTCTTTAGGTAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 420
 DB |||||
 DB 565 GACTCTTTAGGTAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 624
 QY 421 GCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGTCCCA 480
 DB |||||
 DB 625 GCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATACCCAGTAGTCCCA 684
 QY 481 GGGAAAAACGCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
 DB |||||
 DB 685 GGGAAAAACGCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 744
 QY 541 ATAGAGAAAGATGGAATGCGATTAGGATGCTATATTTTTCAGACAGACGCTAGTGGGA 600
 DB |||||
 DB 745 ATAGAGAAAGATGGAATGCGATTAGGATGCTATATTTTTCAGACAGACGCTAGTGGGA 804
 QY 601 GCTGTAGAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTTCAGCTTAAGCAGCA 660
 DB |||||
 DB 805 GCTGTAGAAATGCTAAATCTTAATAACAGTATAGCAACATAGATTTCAGCTTAAGCAGCA 864
 QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGACTCTTCCAATT 720
 DB |||||
 DB 865 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGACTCTTCCAATT 924
 QY 721 CTTCAAGAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 780
 DB |||||
 DB 925 CTTCAAGAGCGGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 984
 QY 781 GCAGATGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAGACACACAGGA 840

505	Db	AAAGTTGGCGGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATATACCAAGCGAATCTTT	564
361	QY	GACTCTTTTAGGTAAACTGACTTCCTCTGGAGCCTCTTTACAGGCTGCTCTTTCTCCAACTCTGTGA	420
565	Db	GACTCTTTTAGGTAAACTGACTTCCTCTGGAGCCTCTTTACAGGCTGCTCTTTCTCCAACTCTGTGA	624
421	QY	GCAACCAATACCAAGCAGCTGAGCTTCCTTTAAAGAGATGCAAGATACCCAGTAGTCCCA	480
625	Db	GCAACCAATACCAAGCAGCTGAGCTTCCTTTAAAGAGATGCAAGATACCCAGTAGTCCCA	684
481	QY	GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG	540
685	Db	GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG	744
541	QY	ATAGAGAAAGATGGAATGCGATTAGGGATGTCATATTTTGAGGACAGAAACGCTAGTGGGA	600
745	Db	ATAGAGAAAGATGGAATGCGATTAGGGATGTCATATTTTGAGGACAGAAACGCTAGTGGGA	804
601	QY	GCTGTGAAAAATGCTAAATCTTAATPAAACAGTATAGCAACATAGATTTCAGCTTAAGACGCA	660
805	Db	GCTGTGAAAAATGCTAAATCTTAATPAAACAGTATAGCAACATAGATTTCAGCTTAAGACGCA	864
661	QY	ATGCTACTGCTAAGACACAATAGCTGAAAGCTCAGAAAAAGTTCCCGCACTCTCCAAATT	720
865	Db	ATGCTACTGCTAAGACACAATAGCTGAAAGCTCAGAAAAAGTTCCCGCACTCTCCAAATT	924
721	QY	CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAAACTT	780
925	Db	CTTCAAGAAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAAACTT	984
781	QY	GCAGATGGTTCTGATGTTTCCAAATCCAGGAACCTACAGTTGAGGAGCTCCAAAGCAACAGGA	840
985	Db	GCAGATGGTTCTGATGTTTCCAAATCCAGGAACCTACAGTTGAGGAGCTCCAAAGCAACAGGA	1044
841	QY	AGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT	900
1045	Db	AGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT	1104
901	QY	TCCATTTTGATGCTCGGGTTTCGTCAGATGTCACATGTTCAATACGGAAAAATCCTGAT	960
1105	Db	TCCATTTTGATGCTCGGGTTTCGTCAGATGTCACATGTTCAATACGGAAAAATCCTGAT	1164
961	QY	TCTCAAGCTGCCCAACAGGAGCTCGACAGCAAAAGCTAGACAGCAAGCGCTGGAGAT	1020
1165	Db	TCTCAAGCTGCCCAACAGGAGCTCGACAGCAAAAGCTAGACAGCAAGCGCTGGAGAT	1224
1021	QY	GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTTAGAAGCGGCTCTAGGTTAA	1080
1225	Db	GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTTAGAAGCGGCTCTAGGTTAA	1284
1081	QY	GCTGGGCAACAAAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG	1140
1285	Db	GCTGGGCAACAAAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG	1344
1141	QY	AGCGCAGGAGTTCCTCCGCTGCAGCAAGTTCATAGGGTCATCTGTAAACAGCTTTAC	1200
1345	Db	AGCGCAGGAGTTCCTCCGCTGCAGCAAGTTCATAGGGTCATCTGTAAACAGCTTTAC	1404
1201	QY	AAGACCTCAAATCTACAGGTTCTGATTTAAAAACACAGATATCAGCAGGTTATGATGCT	1260
1405	Db	AAGACCTCAAATCTACAGGTTCTGATTTAAAAACACAGATATCAGCAGGTTATGATGCT	1464
1261	QY	TACAAATCCATCAATGATGCTATTGGTAGGGCAGAAATGATGGGACTCGTAGTGTGATA	1320
1465	Db	TACAAATCCATCAATGATGCTATTGGTAGGGCAGAAATGATGGGACTCGTAGTGTGATA	1524
1321	QY	AACAATGTAGTACCCCGCTCTCACACGATCCGTTCCCTAGAGCAGCAACAGAGCTCGA	1380
1525	Db	AACAATGTAGTACCCCGCTCTCACACGATCCGTTCCCTAGAGCAGCAACAGAGCTCGA	1584
1381	QY	GGACCGAAAAAAACAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATACAGAACTCTT	1440

RESULT 6

RESOL 6
ART. 91190

ABL91190
ID ABL91190 standard; DNA; 1956 BP.

XX
XX

AC ABL91190;

XX

DT 29-AUG-2003 (revised)

DT 29-JUL-2002 (first entry)

DE Chlamydia pneumoniae cp7033 ORF DNA, SEQ ID NO:14.

[illegible]

KW Chlamydial infection; antigen; immunogen; vaccine; diagnosis;
Chlamydia; antigen; immunogen; vaccine; diagnosis;

KW human respiratory disease; cardiovascular disease; atherosclerosis;

KW coronary artery disease; carotid artery stenosis; myocardial infarction;

KW cerebrovascular disease; aortic aneurysm; claudication; stroke;

strain CWL029; open reading frame; ORF; gene; ds.

XX
30
XX

US Chlamydia pneumoniae pneumonae.
XX

XX	Key	Location/Qualifiers
FH		

FT	key	location/Quadrants	1. 1956
FT	key	location/Quadrants	1. 1956

```

FTT
CDD
T: 1.1330
/*lag= a

```

FT /cag= /product="cp7033"

XX

PN WO200202606-A2.

PD 10-JAN-2002.

PF / 03-JUL-2001; 2001WO-IB001445.

XX

03-JUL-2000; 2000GB-00016363.

PR 11-JUL-2000; 2000GB-00017047.

PR 21-JUL-2000; 2000GB-00017983.

PR 07-AUG-2000; 2000GB-00019368.

18-AUG-2000; 2000GB-00020440.
14-SEP-2000; 2000GB-00033583.
18-AUG-2000; 2000GB-00020440.

14-SEP-2000; 2000GB-00022583.
10-NOV-2000; 2000GB-00037548

PR 10-NOV-2000; 2000GB

FR/XX
22-DEC-2000; 2000GB

PA (CHTR-) CHTRON SPA

XX
FA (CHIR-) CHIRON SPA.

PI Ratti G. Grandi G:

PT Novel Chlamydia pneumoniae protein useful in the manufacture of a
PT medicament for treatment or prevention of infection due to Chlamydia,
XX preferably Chlamydia pneumoniae, and for diagnostic purposes.
XX
PS Claim 5; Page 47-48; 364pp; English.

XX
CC Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia
CC pneumoniae (strain CW029), and ABL91184-ABL91373 represent DNA encoding
CC them. The proteins are predicted to be immunogenic and may therefore be
CC useful in vaccine production and for diagnostic purposes. Chlamydia
CC pneumoniae is a common cause of respiratory disease in humans, and is
CC also involved in the development of cardiovascular diseases such as
CC atherosclerosis, coronary artery disease, carotid artery stenosis,
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
CC claudication and stroke. The proteins and nucleic acids of the invention
CC may be used in vaccines and pharmaceutical compositions for the
CC prevention or treatment of chlamydial infections, particularly Chlamydia
CC pneumoniae infections. The proteins may also be used in the detection of
CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
CC DNA probe assay or blotting techniques for determining Chlamydia
CC pneumoniae gene expression. The present sequence represents a
CC specifically claimed DNA which encodes a Chlamydia pneumoniae protein of
CC the invention. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 1955 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;

Query Match 34.4%; Score 1748.8; DB 6; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTACGAGA	60
DB	205	ATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTACTACGAGA	264
QY	61	TCTGACAGCTGGACTCAACGACAGCAGCAGCCACCTACGCTCTCTCCACCCAGCTTTGAT	120
DB	265	TCTGACAGCTGGACTCAACGACAGCAGCAGCCACCTACGCTCTCTCCACCCAGCTTTGAT	324
QY	121	GATTATAGACTCAGCGGCAACAGCTTAGATACTATCTTTACCTCAACATCACTAGCT	180
DB	325	GATTATAGACTCAGCGGCAACAGCTTAGATACTATCTTTACCTCAACATCACTAGCT	384
QY	181	GACATACAGGCTGCTTGGTGAGCTCCAGGATGCTCTACTAATATAAAGGATACAGCG	240
DB	385	GACATACAGGCTGCTTGGTGAGCTCCAGGATGCTCTACTAATATAAAGGATACAGCG	444
QY	241	GCTACTGATGAGGAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCGATT	300
DB	445	GCTACTGATGAGGAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCGATT	504
QY	301	AAAGTTGGCGGCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGGATCTTT	360
DB	505	AAAGTTGGCGGCAAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGGATCTTT	564
QY	361	GACTCTTTAGGTAACAGTACTCTCTTGACCTCTTACAGGCTGCTCTTCTCCCAATCTGTA	420
DB	565	GACTCTTTAGGTAACAGTACTCTCTTGACCTCTTACAGGCTGCTCTTCTCCCAATCTGTA	624
QY	421	GCAACAAATAAACAGCAGCTGAGCTTCTTAAAGAGATGAAGAAATCCAGTAGTCCCA	480
DB	625	GCAACAAATAAACAGCAGCTGAGCTTCTTAAAGAGATGAAGAAATCCAGTAGTCCCA	684
QY	481	GGGAAACCGCTGCAATGCTCAATCTTTAGTTAGTACAGACAGATGCTACAGGACACAG	540
DB	685	GGGAAACCGCTGCAATGCTCAATCTTTAGTTAGTACAGACAGATGCTACAGGACACAG	744
QY	541	ATAGAGAAAGATGAAATGCGATTAGGATGCAATATTTTCAGGACAGAAACGCTAGTGA	600
DB	745	ATAGAGAAAGATGAAATGCGATTAGGATGCAATATTTTCAGGACAGAAACGCTAGTGA	804
QY	601	GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTAGTAAAGCAGCA	660
DB	805	GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTAGTAAAGCAGCA	864

QY	661	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT	720
DB	865	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT	924
QY	721	CTTCAAGAGCGGACAAATGGTATATACAGGCTGAGAAAGATCTTAAAATATCAAACCT	780
DB	925	CTTCAAGAGCGGACAAATGGTATATACAGGCTGAGAAAGATCTTAAAATATCAAACCT	884
QY	781	GCAGATGGTCTGATCTTCCAAATCCAGGAACTACAGTTTGGAGCTCCAAAGCAACAGGA	840
DB	985	GCAGATGGTCTGATCTTCCAAATCCAGGAACTACAGTTTGGAGCTCCAAAGCAACAGGA	1044
QY	841	AGTAGTATGGTAGTATTCGTGTTCCATGCTGTTTAGATGATGCTGAAAATAGACCGCT	900
DB	1045	AGTAGTATGGTAGTATTCGTGTTCCATGCTGTTTAGATGATGCTGAAAATAGACCGCT	1104
QY	901	TCCATTTTATGCTCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT	960
DB	1105	TCCATTTTATGCTCTGGGTTTCTGTCAGATGATTCACATGTTCAATACGGAATCTCTGAT	1164
QY	961	TCTCAAGCTGCCCAACAGGAGCTCGCAGCACCAAGCTTAGAGCAGCGAAAGCCGCTGAGAT	1020
DB	1165	TCTCAAGCTGCCCAACAGGAGCTCGCAGCACCAAGCTTAGAGCAGCGAAAGCCGCTGAGAT	1224
QY	1021	GACAGTCTGCTGACGCTGGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGGTAAA	1080
DB	1225	GACAGTCTGCTGACGCTGGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGGTAAA	1284
QY	1081	GCTGGSCAACACAGGSCATCTCAATGCTTTAGGACAGATCGCTCTGCTGCTGTG	1140
DB	1285	GCTGGSCAACACAGGSCATCTCAATGCTTTAGGACAGATCGCTCTGCTGCTGTG	1344
QY	1141	AGCGCAGGAGTTCTCTCCGCTCGCAGCAAGTCTTAGGGTCTATCTGTAAAACAGCTTTAC	1200
DB	1345	AGCGCAGGAGTTCTCTCCGCTCGCAGCAAGTCTTAGGGTCTATCTGTAAAACAGCTTTAC	1404
QY	1201	AGACCTCAAAATCTACAGGTTCTGATATAAAACACAGATATCAGCAGGTTATGATGT	1260
DB	1405	AGACCTCAAAATCTACAGGTTCTGATATAAAACACAGATATCAGCAGGTTATGATGT	1464
QY	1261	TACAAATCCATCAATGATGCTTATGGTGGCAGCAAGTATGATGCGACTCGTATGTGATA	1320
DB	1465	TACAAATCCATCAATGATGCTTATGGTGGCAGCAAGTATGATGCGACTCGTATGTGATA	1524
QY	1321	AACAAATGTAAGTACCCCGCTCTCACACATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA	1380
DB	1525	AACAAATGTAAGTACCCCGCTCTCACACATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA	1584
QY	1381	GGACCAAGAAAACAGATCAAGCCCTCGTAGGTGATTTCTGGCAATAGCAGAACTCTT	1440
DB	1585	GGACCAAGAAAACAGATCAAGCCCTCGTAGGTGATTTCTGGCAATAGCAGAACTCTT	1644
QY	1441	GGAGATGCTCTAGTCAAGTTTCGGCACTACCAATCTGTAATCGCATCACTCAGTCGAAT	1500
DB	1645	GGAGATGCTCTAGTCAAGTTTCGGCACTACCAATCTGTAATCGCATCACTCAGTCGAAT	1704
QY	1501	CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAGGCTTCA	1560
DB	1705	CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAGGCTTCA	1764
QY	1561	CAGTTTGGGTATCTTATGTCGAACTTTCTATAGTCTTACACAGAGTTCTATAGTAAA	1620
DB	1765	CAGTTTGGGTATCTTATGTCGAACTTTCTATAGTCTTACACAGAGTTCTATAGTAAA	1824
QY	1621	TTAGAAAGTTTCTTGTGCTGAAGGATCTAGGACAGAGCTGTAATAAAGCACTTTTCTTT	1680
DB	1825	TTAGAAAGTTTCTTGTGCTGAAGGATCTAGGACAGAGCTGTAATAAAGCACTTTTCTTT	1884
QY	1681	GAAACGAACTCTTGTATTCAGCAGGTCGTGGTCAATATCGGCTCTCTATATTTCTGT	1740
DB	1885	GAAACGAACTCTTGTATTCAGCAGGTCGTGGTCAATATCGGCTCTCTATATTTCTGT	1944

QY	1741	TATCTCCATATA	1752	QY	1	ATGAGCTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	60
Db	1945	TATCTCCATATA	1956	Db	205	ATGAGCTCTGGCAGATAAGCTGGGTATTCCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	264
RESULT 7				QY	61	TCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCTCCCTCCACCCACGTTTGAT	120
ADD42825				Db	265	TCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCTCCCTCCACCCACGTTTGAT	324
ID ADD42825 standard; DNA; 1956 BP.				QY	121	GATTATAAGACTCAAGCGCAAAACAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT	180
XX ADD42825;				Db	325	GATTATAAGACTCAAGCGCAAAACAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT	384
XX 15-JAN-2004 (first entry)				QY	181	GACATACAGGCTGCTTTGGTGAGCCCTCCAGAGTCTGTCTCACTAATATAAGGATACAGCG	240
XX Chlamydia pneumoniae antigen polynucleotide SEQ ID NO:63.				Db	385	GACATACAGGCTGCTTTGGTGAGCCCTCCAGAGTCTGTCTCACTAATATAAGGATACAGCG	444
XX Chlamydia infection; Chlamydia; antibiotic; antiinflammatory;				QY	241	GCTACTGATGAGAAACCGCAATCCCTCGGAGTGGGAACTTAAGAACTTAAGAACTTAAGAA	300
XX antiinfectivity; cardiant; antiarteriosclerotic; ophthalmological;				Db	445	GCTACTGATGAGAAACCGCAATCCCTCGGAGTGGGAACTTAAGAACTTAAGAACTTAAGAA	504
XX vaccine; gene therapy; immune response; pelvic inflammatory disease;				QY	301	AAAGTTGGCGCGCAAAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAA	360
XX tubal obstruction; infertility; male infertility; ocular infection;				Db	505	AAAGTTGGCGCGCAAAATTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAACTTAAGAA	564
XX blindness; acute respiratory tract infection; atherosclerosis;				QY	361	GACTCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA	420
XX coronary heart disease; gene; ds.				Db	565	GACTCTTTAGGTAAACTGACTTCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA	624
XX Chlamydia pneumoniae.				QY	421	GCATAAATAAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAGTCCCA	480
OS WO2003041560-A2.				Db	625	GCATAAATAAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCAGTAGTCCCA	684
XX 22-MAY-2003.				QY	481	GGGAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAG	540
XX 05-NOV-2002; 2002WO-US035624.				Db	685	GGGAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAG	744
XX 06-NOV-2001; 2001US-00012256.				QY	541	ATAGAGAAAGATGGAATGCGATAGGATGAGTATTTTTCAGGACAGAAACGCTAGTGA	600
XX 05-DEC-2001; 2001US-00007693.				Db	745	ATAGAGAAAGATGGAATGCGATAGGATGAGTATTTTTCAGGACAGAAACGCTAGTGA	804
XX 15-JUL-2002; 2002US-00197220.				QY	601	GCTGTAGAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTTAAGACGA	660
XX (CORI-) CORIXA CORP.				Db	805	GCTGTAGAAATGCTTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTTAAGACGA	864
XX Bhatia A, Guiderian J, Skeiky YAW, Maisonneuve JL, Barth B;				QY	661	ATCGTACTGCTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT	720
XX Probst P;				Db	865	ATCGTACTGCTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT	924
XX WPI: 2003-441771/41.				QY	721	CTTCAAGAGCGGAAACAAATGCTTAATACAGGCTCAGAAAGATCTTAAATAATCAACCT	780
XX P-PSDB; ADD42865.				Db	925	CTTCAAGAGCGGAAACAAATGCTTAATACAGGCTCAGAAAGATCTTAAATAATCAACCT	984
XX New DNA and proteins comprising a portion of a Chlamydia antigen, useful				QY	781	GCAGATGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA	840
XX for diagnosing or treating Chlamydia infections, particularly as				Db	985	GCAGATGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA	1044
XX vaccines for treating or preventing Chlamydia infections, e.g. pelvic				QY	841	AGTAGTATTGTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT	900
XX inflammatory disease.				Db	1045	AGTAGTATTGTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT	1104
XX Disclosure; SEQ ID NO 63; 275pp; English.				QY	901	TCCATTTTGTATGCTGTTTTCGTTTCGTCAGATGATTCACATGTTCAATCGGAAATCTTGAT	960
XX The present invention describes compounds and methods for diagnosing and				Db	1105	TCCATTTTGTATGCTGTTTTCGTTTCGTCAGATGATTCACATGTTCAATCGGAAATCTTGAT	1164
XX treating Chlamydia infection. Chlamydia polynucleotide and protein				QY	961	TCTCAAGTGTCCCAACAGGAGCTCGCAGCACAAGCTAGACAGCAAGCGGCTGAGAT	1020
XX sequences have antibiotic, antiinflammatory, antiinfectivity, cardiant,				Db	1165	TCTCAAGTGTCCCAACAGGAGCTCGCAGCACAAGCTAGACAGCAAGCGGCTGAGAT	1224
XX antiarteriosclerotic and ophthalmological activities, and can be used in				QY	1021	GACAGTGTCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTA	1080
XX vaccines, and in gene therapy. The Chlamydia polynucleotides, proteins,				Db	1225	GACAGTGTCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGGTA	1284
XX compositions or methods from the present invention can be used for the				QY	1081	GCTGGCAACACACAGGCGCTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTG	1140
XX serodiagnosis or treatment of Chlamydia infections, particularly in							
XX humans. The polynucleotides, proteins or compositions are particularly							
XX useful for stimulating an immune response in a patient, or for							
XX stimulating and/or expanding T cells specific for a Chlamydia protein.							
XX Specifically, the polynucleotides, proteins or compositions are useful as							
XX vaccines for treating or preventing Chlamydia infections including							
XX pelvic inflammatory disease (which results in tubal obstruction and							
XX infertility in women), male infertility, ocular infection (which may							
XX cause blindness), acute respiratory tract infections, atherosclerosis, or							
XX coronary heart disease. The present sequence is used in the							
XX exemplification of the present invention.							
XX Sequence 1956 BP; 606 A; 436 C; 441 G; 473 T; 0 U; 0 Other;							
Query Match							
Best Local Similarity							
Matches 1750; Conservative							
94.4%; Score 1748.8; DB 9; Length 1956;							
99.9%; Pred. No. 0;							
0; Mismatches							
2; Indels							
0; Gaps							
0;							

Db 1285 GCTGGCAACACAGGGGATCACTCAATGCTTTAGACAGATCGTTCGCTGCTGTGTG 1344
Qy 1141 AGCGCAGAGTTCCTCCGCTGCAGCAAGTTCATAGGCTCATCTGTAACACAGCTTTAC 1200
Db 1345 AGCGCAGAGTTCCTCCGCTGCAGCAAGTTCATAGGCTCATCTGTAACACAGCTTTAC 1404
Qy 1201 AAGACCTCAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
Db 1405 AAGACCTCAAAATCTACAGGTTCTGATTTATAAACAACAGATATCAGCAGGTTATGATGCT 1464
Qy 1261 TACAATCCATCAATGATGCTGATGCTAGGCAAGCAATGATGCTGCTGATGCTGATA 1320
Db 1465 TACAATCCATCAATGATGCTGATGCTAGGCAAGCAATGATGCTGCTGATGCTGATA 1524
Qy 1321 AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGACACAGCAAGAGCTCGA 1380
Db 1525 AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGACACAGCAAGAGCTCGA 1584
Qy 1381 GGACAGAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGACTCTT 1440
Db 1585 GGACAGAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGACTCTT 1644
Qy 1441 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCAGATCACTCAGTCCGAAT 1500
Db 1645 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAATGCAGATCACTCAGTCCGAAT 1704
Qy 1501 CCTCAAGCAATATAGAGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCTTCCA 1560
Db 1705 CCTCAAGCAATATAGAGAGATCAGACAAAGCTTACATCGGAGTGCACAAAGCTTCCA 1764
Qy 1561 CAGTTGGCTATCCTTATGCTCACTTCTATGACTCTACACAGAGTTCATAGCTAAA 1620
Db 1765 CAGTTGGCTATCCTTATGCTCACTTCTATGACTCTACACAGAGTTCATAGCTAAA 1824
Qy 1621 TTAGAAAGTTGTTTGCTGAAGGATCTAGGACAGCAGCTGAAATATAAAGCACTTTCCTTT 1680
Db 1825 TTAGAAAGTTGTTTGCTGAAGGATCTAGGACAGCAGCTGAAATATAAAGCACTTTCCTTT 1884
Qy 1681 GAACGAACTCTGTTTATTCAGCAGAGTGTGCTCAATTCGGCTCTCTATATCTGGT 1740
Db 1885 GAACGAACTCTGTTTATTCAGCAGAGTGTGCTCAATTCGGCTCTCTATATCTGGT 1944
Qy 1741 TATCTCCAAATA 1752
Db 1945 TATCTCCAAATA 1956

RESULT 8

AA02066

ID AA02066 standard; DNA; 2238 BP.

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

XX AC AA02066;

FT misc_feature /product= "Truncated Chlamydia pneumoniae 76KDa protein"
FT 2122..2238
FT /*tag= C
FT /note= "This part of the sequence is unrelated to C.
FT pneumoniae 76 kDa gene"
XX W020066739-A2.
XX 09-NOV-2000.
XX 03-MAY-2000; 2000WO-CA000511.
XX 03-MAY-1999; 99US-0132270P.
XX 30-JUN-1999; 99US-0141276P.
XX (AVET) AVENTIS PASTEUR LTD.
XX Murdin AD, Oomen RP, Wang J, Dunn P;
XX WPI; 2000-687542/67.
XX P-PSDB; AAY1957.
XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful
XX for vaccinating against Chlamydia infections.
XX Claim 32; Fig 3; 112pp; English.
XX The present sequence is a DNA coding for a fusion protein comprising a
XX truncated Chlamydia pneumoniae 76 kDa protein and vector-encoded
XX residues. C. pneumoniae 76 kDa protein is used in the diagnosis,
XX prevention and treatment of C. pneumoniae infections (e.g. pneumonia,
XX upper respiratory tract disease, bronchitis, sinusitis and acute
XX respiratory disease such as cough, sore throat, hoarseness, fever; and
XX abnormal chest sounds on auscultation). C. pneumoniae sequence is also
XX used as vaccines for immunising humans against diseases caused by C.
XX pneumoniae
XX Sequence 2238 BP; 689 A; 459 C; 497 G; 593 T; 0 U; 0 Other;
Query Match 62.7%; Score 1161.8; DB 3; Length 2238;
Best Local Similarity 99.7%; Pred. No. 8.3e-314;
Matches 1174; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 ATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
Db 970 ATGAGTCTGGCAGATAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 1029
Qy 61 TCTGCAGACGTGGACTCAACGACGAGCGACCGCTACGCTCTCTCCACCCAGCTTTGAT 120
Db 1030 TCTGCAGACGTGGACTCAACGACGAGCGACCGCTACGCTCTCTCCACCCAGCTTTGAT 1089
Qy 121 GATTATAAGACTCAAGCGAAACAGCTTAGTACTACTTCTTACCTCAACATCACTAGCT 180
Db 1090 GATTATAAGACTCAAGCGAAACAGCTTAGTACTACTTCTTACCTCAACATCACTAGCT 1149
Qy 181 GACATACAGCTGTTTGGTGAGCTCCAGATGCTCTCACTAATAATAAGGATACAGCG 240
Db 1150 GACATACAGCTGTTTGGTGAGCTCCAGATGCTCTCACTAATAATAAGGATACAGCG 1209
Qy 241 GCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACCTAAGATGCCGATGAGTT 300
Db 1210 GCTACTGATGAGGAAACCGCAATCGCTGCGAGTGGGAAACCTAAGATGCCGATGAGTT 1269
Qy 301 AAGATTGGCGCGCAAAATTACAGATTAAGCAATATGCTTCGGATTAACCAAGGATCTTT 360
Db 1270 AAGATTGGCGCGCAAAATTACAGATTAAGCAATATGCTTCGGATTAACCAAGGATCTTT 1329
Qy 361 GACTCTTTAGTAACTGACTTCTCTCGACTCTTACAGGCTGCTCTTCTCCCAATCTGTA 420
Db 1330 GACTCTTTAGTAACTGACTTCTCTCGACTCTTACAGGCTGCTCTTCTCCCAATCTGTA 1389
Qy 421 GCAACAATAAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATTAACCCAGTAGTCCCA 480

Db 1390 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 1449
Qy 481 GGGAAAAAGCCCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAG 540
Db 1450 GGGAAAAAGCCCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAG 1509
Qy 541 ATAGAGAAAGATGGAATGCGAATAGGATGCGATATTTGACAGGACAGACCTAGTGG 600
Db 1510 ATAGAGAAAGATGGAATGCGAATAGGATGCGATATTTGACAGGACAGACCTAGTGG 1569
Qy 601 GCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTAAAGCAGCA 660
Db 1570 GCTGTAGAAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTAAAGCAGCA 1629
Qy 661 ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCTCAATT 720
Db 1630 ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCTCAATT 1689
Qy 721 CTTCAAGAAAGCGGAACAAATGCTAATACAGGCTGAGAAAGATCTTTAAAAATATCAAACT 780
Db 1690 CTTCAAGAAAGCGGAACAAATGCTAATACAGGCTGAGAAAGATCTTTAAAAATATCAAACT 1749
Qy 781 GCAGATGCTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAAGGA 840
Db 1750 GCAGATGCTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAAGGA 1809
Qy 841 AGTAGTATGCTAGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCCCT 900
Db 1810 AGTAGTATGCTAGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCCCT 1869
Qy 901 TCCATTTTGTGCTGCTGGTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 960
Db 1870 TCCATTTTGTGCTGCTGGTTCGTCAGATGATTCACATGTTCAATACGGAATAATCCTGAT 1929
Qy 961 TCTCAAGCTGCCAACAGAGAGCTGCAGACACAAGCTAGACAGCGAAAGCCGCTGGAGAT 1020
Db 1930 TCTCAAGCTGCCAACAGAGAGCTGCAGACACAAGCTAGACAGCGAAAGCCGCTGGAGAT 1989
Qy 1021 GACAGTGTCTGTCAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
Db 1990 GACAGTGTCTGTCAGCGCTGCGAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 2049
Qy 1081 GCTGGGCAACACAGGGGATCTCAATGCTTTAGACAGATGCTGCTGCTGCTGTTG 1140
Db 2050 GCTGGGCAACACAGGGGATCTCAATGCTTTAGACAGATGCTGCTGCTGCTGTTG 2109
Qy 1141 AGCGCAGAGTTCCTCCCGCTGCGACAGATTCATAG 1177
Db 2110 AGCGCAGAGTTCCTCCCGCTGCGACAGATTCATAG 2145

RESULT 9

AD02065
ID AAD02065 standard; DNA; 1456 BP.
AC AAD02065;
DT 26-MAR-2001 (first entry)
DE 3'-truncated Chlamydia pneumoniae 76 kDa protein encoding DNA.
XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
KW upper respiratory tract disease; bronchitis; sinusitis;
KW acute respiratory disease; cough; sore throat; hoarseness; fever;
KW vaccine; immunisation; treatment; truncation mutant; ds.
XX Chlamydia pneumoniae.
OS Synthetic.

XX Key Location/Qualifiers
FH 101..1456
CDS /*tag= a
FT /product= "3'-truncated Chlamydia pneumoniae 76kDa
FT

protein"
/note= "The coding region does not include stop codon"
/partial

W0200066739-A2.

XX 09-NOV-2000.

XX 03-MAY-2000; 2000WO-CA000511.

XX 03-MAY-1999; 99US-0132270P.

XX 30-JUN-1999; 99US-0141276P.

XX (AVET) AVENTIS PASTEUR LTD.

XX Murdin AD, Oomen RP, Wang J, Dunn P;

XX WPI; 2000-687542/67.

XX P-PSDB; AAY71956.

Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful for vaccinating against Chlamydia infections.

XX Claim 2c; Page 102-104; 11:2pp; English.

The present sequence is a DNA coding for 3'-truncated Chlamydia pneumoniae 76 kDa protein. C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory tract disease, bronchitis, sinusitis and acute respiratory disease such as cough, sore throat, hoarseness, fever; and abnormal chest sounds on auscultation). C. pneumoniae sequence is also used as vaccines for immunising humans against diseases caused by C. pneumoniae

XX Sequence 1456 BP; 452 A; 308 C; 331 G; 365 T; 0 U; 0 Other;

Query Match 62.18; Score 1151; DB 3; Length 1456;

Best Local Similarity 100.0%; Pred. No. 6.6e-311;

Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 60
Db 305 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTGCTCTTCTACTAGCAGA 364
Qy 61 TCTGCAGACGTGGACTCAACAGACAGCCGACCTACGCTCCTCCACCCACGTTTGTAT 120
Db 365 TCTGCAGACGTGGACTCAACAGACAGCCGACCTACGCTCCTCCACCCACGTTTGTAT 424
Qy 121 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTACTCAACATCACTAGCT 180
Db 425 GATTATAAGACTCAAGCGCAACAGCTTACGATATCTTTACTCAACATCACTAGCT 484
Qy 181 GACATACAGCTGCTTTTGGTGAGCTTCCAGGATGCTCTCACTAATAAAGGATACAGCG 240
Db 485 GACATACAGCTGCTTTTGGTGAGCTTCCAGGATGCTCTCACTAATAAAGGATACAGCG 544
Qy 241 GCTACTGATGAGGAACCGCAATCGCTGCGAGTGGGAAACTAAGATGGCGATG 300
Db 545 GCTACTGATGAGGAACCGCAATCGCTGCGAGTGGGAAACTAAGATGGCGATG 604
Qy 301 AAAGTTGGCGGCAATTTACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTCTT 360
Db 605 AAAGTTGGCGGCAATTTACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTCTT 664
Qy 361 GACTCTTTAGGTAAACTGACTTCTCTTGACCTCTTACAGGCTGCTCTTCTCCAACTGTGA 420
Db 665 GACTCTTTAGGTAAACTGACTTCTCTTGACCTCTTACAGGCTGCTCTTCTCCAACTGTGA 724
Qy 421 GCAAACAAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db 725 GCAAACAAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 784
Qy 481 GGGAAACGCTGCAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACAG 540

WP	AAZ01425_07	700001	810000	
WP	AAZ01425_08	800001	910000	
WP	AAZ01425_09	900001	1010000	
WP	AAZ01425_10	1000001	1038602	

Query Match		10.1%;	Score 187.6;	DB 2;	Length 110000;
Best Local Similarity		53.0%;	Pred. No. 1.2e-40;		
Matches	480;	Conservative	0;	Mismatches 408;	Indels 18; Gaps 3;
QY	863	TTTCCATGCTGTTAGATGATGCTGAAATAGACGGCTTCCATTTTCATTTTCATGTCGGGTTTC	922		
DB	5378	TTTCTTTGTTGTTGATGATGACAAATGAAATGCGACGATTCGAATGCAAGTTTTC	5437		
QY	923	GTCAATGATTCACATGTTTCAATACGGAAATCTCGATTTCAAGCTGCCAACACAGAGC	982		
DB	5438	GATCTATGATCGAAACAATTTAATCTAAACAATCTGCAACAGCTTAAAGAGCTACAGCTA	5497		
QY	983	TCGCAGCACAAAGCTPAGAGCAGCGAAAGCCGTGGAGATGACAGTGTCTGTCAGCGCTGG	1042		
DB	5498	TGGAGGCTCAGCTGA---CTGCGATGTCAATCACTGGTTGGTGGCGGATGGGAGCTCC	5554		
QY	1043	CAGATGCTCAGAAAGCTTTGAAGCGGCTCTAGTAAAGCTGGGCAACAAC-----	1093		
DB	5555	CAGCGAAATACAAGCAATCAAAGATGCTCTTGGCGAAGCTTTGAAACAACCATCAGCAG	5614		
QY	1094	AGGCACTACTCAATGCTTTTAGGACAGATTCGCTTCTGCTGCTGTGTGTCAGCGCAGGATTC	1153		
DB	5615	ATGGTTTAGCTACAGCTATGGGACAAGTGGCTTTTGCAAGCTGCCAAGGTTGGAGGAGCT	5674		
QY	1154	CTCCCGCTCAGCAAGTTCTATAGGTCATCTGTAAACAAGCTTTTCAAGACCTCAAAAT	1213		
DB	5675	CCGCAGAAACAGCTGGGCACTGTCCAGATGAATGTAAACAAGCTTTTCAAGACAGCGTTT	5734		
QY	1214	CTACAGGTTCT-----GATTATAAAACACAGATATCAGCAGGTTATGATGCTTACAAAT	1267		
DB	5735	CTTCGACTTCTTCAAGCTCTTATGCAGCAGCACCTTCCGATGGATATTCGCTTACAAA	5794		
QY	1268	CCATCAATGATGCTATGTTAGGCACGAATGATGCGACTCGTGATGTGATTAACATG	1327		
DB	5795	CATGTAACCTTTTATATTCGAAAGCAGAGCGCGTGCAGTCAGCTATTTAGTCAAACTG	5854		
QY	1328	TAAAGTACCCCGCTCTCACAGATCCGTTCCCTAGAGCACGAACAGAGCTCGAGGACCAG	1387		
DB	5855	CAAAATCCGCGCTTCCAGAGCGTTTCTCGTTCTGGCATAGAAAGTCACAGGACGCACTG	5914		
QY	1388	AAAAACACAGATCAAGCCCTCGTAGGTGATTTCTGGCAATAGCAGAACTCTTTGGAGATG	1447		
DB	5915	CAGATGTAGCCAAAGACGACGAGAACTATTTGCAGAGATAGCCAAACGTTAGGTGATG	5974		
QY	1448	TCTATAGTCAAGTTTCGGCACTCAATCTGTATGCMAGATCATCTAGTCGAATCTCTCAAG	1507		
DB	5975	TATATAGCCGCTTACAGGTTCTGGATTCCTTTGATGTCTACGATGTGAGCAATCCGCAAG	6034		
QY	1508	CGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCAGTGACAAAGCCCTCCACAGTTTG	1567		
DB	6035	CAAAATCAAGAGAGATTTATCGAGAGCTCAGGCACTATTATAGCAAGCTTCCACAAATTG	6094		
QY	1568	GCTATCCTTATGCAACTTTCTTAATGACTCTACACAGAAAGTTTCATAGCTTAAATTAGAAA	1627		
DB	6095	GGTATCTGTGTTCAGAATCTGTGGATAGCTTGCAGAGTTTGTGTCACAAATTGGAAA	6154		
QY	1628	GTTTGTTGCTGAAGGATCTAGCAGCAGCAGCTGAAATAAAGCACTTCTCTTTGAAACGA	1687		
DB	6155	GAGAGTTTGTGTAGTGGGAACGTTAGTCTCCAGAAATCTCAAGAGAAATGCGTTTAGAAAAC	6214		
QY	1688	ACTCCTTGTTTATTCAGCAGGTCGTGGTCAATATCGGCTCTCTATATTTCTGGTTACTCC	1747		
DB	6215	AGCCCGCTTCATTCACAGGTTGTTGGTAAACHTGCCTCTCTATTCTCTGGTTATCTTT	6274		
QY	1748	AATAAC	1753		
DB	6275	CTTAAC	6280		

RESULT 12	
AAA63292	
ID	AAA63292 standard; DNA; 1537 BP.
XX	
AC	AAA63292;
XX	
DT	02-FEB-2001 (first entry)
XX	
DE	C. trachomatis LGV II clone 12-G3-83 coding sequence.
XX	
KW	Chlamydial infection; sexually transmitted disease;
KW	pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW	trachoma; blindness; acute respiratory tract infection; atherosclerosis;
KW	coronary heart disease; antibacterial; ss.
XX	
OS	Chlamydia trachomatis.
XX	
PN	W0200034483-M2.
XX	
PD	15-JUN-2000.
XX	
PF	08-DEC-1999; 95WO-US029012.
XX	
ER	08-DEC-1998; 98US-00208277.
PR	08-APR-1999; 99US-00288594.
PR	01-OCT-1999; 99US-00410568.
PR	22-OCT-1999; 99US-00426571.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PT	Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX	
DR	WPI; 2000-431303/37.
XX	
PT	Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT	comprises immunogenic portion of Chlamydia antigen, which comprises amino
PT	acid sequence encoded by polynucleotide sequence.
XX	
PS	Claim 1; Page 128-129; 256pp; English.
XX	
CC	The present invention relates to new nucleic acid sequences and the
CC	proteins encoded by the nucleic acid sequences. The encoded proteins
CC	comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC	proteins are useful for the serodiagnosis and treatment of Chlamydia
CC	infection. Chlamidiae are intracellular bacterial pathogens that are
CC	responsible for a wide variety of human infections. C. trachomatis
CC	infection is one of the most common sexually transmitted diseases and can
CC	lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC	and infertility. Trachoma due to ocular infection with C. trachomatis is
CC	the leading cause of preventable blindness worldwide. C. pneumonia is a
CC	major cause of acute respiratory tract infections in humans and is also
CC	thought to play a role in the pathogenesis of atherosclerosis and
CC	coronary heart disease. The present sequence is a nucleic acid sequence
CC	isolated in the present invention

Seq	Sequence	1537 BP	459 A	307 C	357 G	414 T	0 U	0 Other
Query	Match	10.1%	Score 187.2	DB 3	Length 1537			
Best Local	Similarity	53.0%	Pred. No. 1.9e-41					
Matches	480	Conservative	0	Mismatches	406	Indels	18	Gaps 3
QY	863	TTTCCATGCTGTTAGATGATGCTGAAATGAGACCGCTTCCATTTTGATGTCGGGTTC	922					
Db	186	TTTCCCTGTTGCTTGATGATGTACATGAAATGGCAGCGATTGCAATGCAAGGTTTTC	245					
QY	923	GTCCAGATGATTCACTGCTTCAATACGGAAATCTCTGATTTCTCAAGCTGCCCAACAGGAGC	982					
Db	246	GATCTATGATCGAACAAATTTAATGTAACAAATCTTCGCAACAGCTAAGAGCTACAAGCTA	305					
QY	983	TCGCAGCACAACTAGACGACGAAAGCCGCTGGAGATGACAGTCTCTGCAGCGCTGG	1042					
Db	306	TGAGGGCTCAGCTGA---CTGCCATGTCAGATCAACTGGTTGGTGGCGATAGCCAGCTCC	362					

RESULT 15

ADD43867
ID ADD43867 standard; DNA; 1944 BP.

XX AC ADD43867;

XX DT 15-JAN-2004 (first entry)

XX DE Chlamydia trachomatis immunogenic DNA sequence, SEQ ID NO 162.

XX KW immunogenic; infection; Chlamydia trachomatis; immune; vaccine;

XX KW gene therapy; antibacterial; gene; ds.

XX OS Chlamydia trachomatis.

XX FN WO2003049762-A2.

XX PD 19-JUN-2003.

XX PF 12-DEC-2002; 2002WO-IB005761.

XX PR 12-DEC-2001; 2001GB-00029732.

XX PR 06-AUG-2002; 2002GB-00018233.

XX PR 14-AUG-2002; 2002GB-00018924.

XX PA (CHIR-) CHIRON SPA.

XX PI Grandi G, Ratti G;

XX DR WPI; 2003-532882/50.

XX DR P-PSDB; ADD43866.

XX PT New immunogenic composition having a protein or encoding nucleic acid,
PT useful for diagnosing, preventing and/or treating Chlamydia trachomatis
PT infection.

XX PS Claim 7; SEQ ID NO 162; 164pp; English.

XX CC The invention relates to a novel immunogenic composition comprising a
CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
CC acid comprises any of 131 fully defined amino acid or nucleotide
CC sequences given in the specification, or has 50% or greater sequence
CC identity to it, or their fragments. The protein and/or nucleic acid of
CC the immunogenic composition is useful in the manufacture of a medicament
CC for the treatment or prevention of infection due to Chlamydia
CC trachomatis. The infection is treated or prevented by the medicament
CC eliciting an immune response which is specific to a C. trachomatis
CC elementary body, or for neutralising C. trachomatis elementary bodies,
CC hence the immunogenic composition can be used in creating a vaccine. The
CC immunogenic compositions can also be used for the diagnosis of C.
CC trachomatis infection. The nucleic acids of the immunogenic compositions
CC can be used to treat disorders by gene therapy. The immunogenic
CC compositions have antibacterial activity. This polynucleotide sequence
CC represents one of the 131 DNA sequences encoding the C. trachomatis
CC proteins with immunogenic properties of the invention.

XX SQ Sequence 1944 BP; 612 A; 391 C; 445 G; 496 T; 0 U; 0 Other;

Query Match 10.0%; Score 186; DB 9; Length 1944;

Best Local Similarity 48.5%; Pred. No. 4.6e-41;

Matches 720; Conservative 0; Mismatches 725; Indels 39; Gaps 6;

QY 302 AAGTTGGCGCGCAATTACAGAAATAGCGAATATGCTTCGGATAACCAAGCATCTCTG 361

DB 467 AGGTTATTGATAGCTCATCTAGTTAGTTAACTTGAATAACAGAAATCAGACTTTAAAGG 526

QY 362 ACTCTTTAGGTTAACTGACTTCTCTTCGACCTTTACAGGCTGCTCTTCTCCAATCTGTAG 421

DB 527 AAACCTTTAACACCAACAGACTCTGCAGATCAGATTCAGGCGATTAAATAGTCAGTTAGAGA 586

QY 422 CAAACAATAACAAACGACTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCGAG 481

DB 587 TCAACAAAAATTCTGCAGATCAAAATTATCAAGATCTGGAAGGCAAAAACATAAGTTATG 646

QY 482 GGAACACGCCCTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTTACGCGACACAGA 541

DB 647 AAGCTGTTCTCACTAACGCGAGGAGGTTATCAAGGCTTCTTCTGGAAGCGGAATTAAAGT 706

QY 542 TAGAGAAAGATGGAAATGCGATTAGGGATGCATATTTTTCAGGACAGAAAGCTAGTGGAG 601

DB 707 TAGACAAAGCTTTGCACTGCTATTGTGATCTGGGATCAAGCGCGGCTGCACTTTCTTC 766

QY 602 CTGTAGAAAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTAAACAGCAA 661

DB 767 AAGCACAGCAAAATAATAGCCAGATAATTCGACGCCACGAGAATAATTATGATGCTG 826

QY 662 TCGCTACTGTCTAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGCAGCTCTTCCAATTC 721

DB 827 CTGAACGAAAGGTTAAACGAGTTAAAAACAAGACATACAGGCTTAACGCACTTCGCTTTAG 886

QY 722 TTCAGAGCGGACAAATGGTATATACAGGCTGGAAGATCTTAAATAATATCAAACTG 781

DB 887 TGAATAAAGCTGAGGACGAGATTAGTCAAGCAAAAAGATAATTCAAGAGATCAAACT- 945

QY 782 CAGATGTTCTGTATGTTTCCAAATCCAGAACTACAGTTGGAG-----GCTCCAAGCAACA 836

DB 946 --AGTGGTTCGGATATTTCTTATCGTTGGTCCGAGTGGTTCAGCTGCTCCGCGAAGTG 1003

QY 837 AGGAAGTAGTATTGGTAGTAT-----TCGTGTTTCCATGCTGTTAGATGATG 893

DB 1004 CGGTAGGAGGCTTGAATCTCTTAACAATTCAAGAAAGAAATTTCCCTTGTCTTGATGATG 1063

QY 884 CTGAAAAATGAGACCGCTTCCATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 943

DB 1064 TAGCAATGAATGGCAGCGATTCGATGCAAGGTTTTCGATCTATGATGCAACAATTA 1123

QY 944 ATACGGAAAAATCTGTGATTTCTCAAGCTCCCAACAGGAGCTCGCAGCAACAAGCTAGACAG 1003

DB 1124 ATGTAAACAATCTCTGCAACAGCTAAAGAGCTAAAGCTATGAGGCTCAGCTGA---CTG 1180

QY 1004 CGAAGCGCTGGAGATGACAGTCTCTGTCAGCGCTGGGAGATGCTCAGAAAGCTTTAG 1063

DB 1181 CGATGTGAGATCAACTGGTTGGTGGGATCGGAGCTCCGACGCGAAATACAAAGCAATCA 1240

QY 1064 AAGCGGCTCTAGTAAAGCTGGGCAACAAC-----AGGGCATACTCAATGCTTTAG 1114

DB 1241 AAGATGCTCTTGGCAAGCTTTGAAACAACCAATCAACAGATGTTTATAGCTACAGCTATGG 1300

QY 1115 GACGATGCGTCTCTGCTGCTGTTGTGAGCGCAGAGTTCCTCCGCTGCGAGCAAGTCTTA 1174

DB 1301 GACAAGTGGCTTTTTCGAGCTGCCAAGGTTGGAGGAGGCTCCGACGAGCAACAGCTGSCACTG 1360

QY 1175 TAGGTCATCTGTAAAAACAGCTTTTACAGACCTCAAAATCTACAGGTTCTGA-----TT 1228

DB 1361 TCCAGATGAATGTAACACAGCTTTTCAAGCAGGTTTTCGACTCTTCCAGCTCTT 1420

QY 1229 ATAAAAACAGATATACGAGGTTTATGATCTTACAAATCCATCAATGATGCTCCATGGTA 1288

DB 1421 ATGCAGCAGCACTTTCCGATGGATATTCTGCTTACAAAAACACTGAACCTTTTATATCCG 1480

QY 1289 GGGCACAAATGATGCGACTCGTGATGTGATAACAATTAAGTAAGTACCCCGCTCTACAC 1348

DB 1481 AAAGCAGAGCGCGGTGCGACTGATTAAGTCAAACTGCAAAATCCCGGCTTTTCCAGAA 1540

QY 1349 GATCCGTTCTTAGAGCAGCAACAGAGCTCGAGACCAGAAAAACAGATCAAGACCCCTCG 1408

DB 1541 GCGTTTCTCGTTCTGGCATAGAAGTCAAGAGCGAGTGCAGATGCTAGCCAAAGACAG 1600

QY 1409 CTAGGGTGATTTCTGGCAATAGCAGAACTCTTTCGAGATGCTTATAGTCAAGTTTTCGSCAC 1468

DB 1601 CAGAAACTATTGTTCAGAGATAGCCAAACGTTAGTGTGATGATATATAGCCGCTTACAGGTT 1660

QY 1469 TACAATCTGTANTGCAGATCACTCAGTCGAATCTTCAGCGCAATTAATGAGAGATCAGAC 1528

DB 1661 TGGATTCTTTGATGCTACGATTGTGAGCAATCCGCAAGTAAATCAAGAGAGATTATGC 1720

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd..

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:23:54 ; Search time 639.252 Seconds

(without alignments)
10716.954 Million cell updates/sec

Title: US-10-608-559-3

Perfect score: 1852
Sequence: 1 atgagctcggagataagct.....gatcttcataataaaaga 1852

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2438257 seqs, 1849576744 residues

Total number of hits satisfying chosen parameters: 4876514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1852	100.0	1852	US-09-564-479-3	Sequence 3, Appli
2	1848.8	99.8	2156	US-09-564-479-1	Sequence 1, Appli
3	1836.2	99.1	1230025	US-10-289-762-1	Sequence 1, Appli
4	1748.8	94.4	1956	US-09-841-132-385	Sequence 385, Appl
5	1748.8	94.4	1956	US-09-841-260-63	Sequence 63, Appl
6	1748.8	94.4	1956	US-10-007-693-63	Sequence 63, Appl
7	1748.8	94.4	1956	US-10-312-273-14	Sequence 14, Appli
8	1161.8	62.7	2238	US-09-564-479-7	Sequence 7, Appli
9	1151	62.1	1456	US-09-564-479-5	Sequence 5, Appli
10	187.2	10.1	1637	US-09-841-132-57	Sequence 57, Appl
11	185.4	10.0	1941	US-09-841-132-412	Sequence 412, Appl
12	184	9.9	1171	US-09-841-260-13	Sequence 13, Appl
13	184	9.9	1171	US-10-007-693-13	Sequence 13, Appl
14	184	9.9	1834	US-09-841-260-38	Sequence 38, Appl
15	184	9.9	1834	US-10-007-693-38	Sequence 38, Appl

16	183	9.9	1983	10	US-09-841-260-28	Sequence 28, Appl
17	183	9.9	1983	13	US-10-007-693-28	Sequence 28, Appl
18	49	2.6	1860	12	US-10-282-122A-37518	Sequence 37518, A
19	47	2.5	9373	12	US-10-221-613-178	Sequence 178, App
20	47	2.5	9373	14	US-10-311-455-1264	Sequence 1264, Ap
21	45.2	2.4	2404	15	US-10-104-047-1740	Sequence 1740, Ap
22	45.2	2.4	3439	15	US-10-161-327-53	Sequence 53, Appl
23	44.4	2.4	2499	9	US-09-842-552-92	Sequence 92, Appl
24	44.2	2.4	4997	12	US-10-282-122A-35506	Sequence 35506, A
25	43.8	2.4	702	15	US-10-037-632-45880	Sequence 45880, A
26	43.8	2.4	702	15	US-10-027-632-45881	Sequence 45881, A
27	42.6	2.3	1431	12	US-10-424-599-127608	Sequence 127608, A
28	42.2	2.3	897	12	US-10-424-599-127609	Sequence 127609, A
29	41.4	2.2	1398	12	US-10-282-122A-15784	Sequence 15784, A
30	41.2	2.2	1237	12	US-10-412-699B-453	Sequence 453, App
31	41.2	2.2	1237	14	US-10-295-403-49	Sequence 49, Appl
32	41.2	2.2	1368	12	US-10-425-114-14565	Sequence 14565, A
33	41.2	2.2	1737	14	US-10-032-585-6793	Sequence 6793, Ap
34	40.6	2.2	618	15	US-10-027-632-269245	Sequence 269245, A
35	40.6	2.2	3249	12	US-10-424-599-139567	Sequence 139567, A
36	40.4	2.2	512	9	US-09-864-761-25347	Sequence 25347, A
37	40.4	2.2	575	9	US-09-864-761-8628	Sequence 8628, Ap
38	40.4	2.2	3137	14	US-10-037-270-714	Sequence 714, App
39	40.4	2.2	3137	15	US-10-117-722-714	Sequence 714, App
40	40.4	2.2	12541	15	US-10-074-024-524	Sequence 524, App
41	40.4	2.2	12541	15	US-10-074-024-525	Sequence 525, App
42	40	2.2	7311	9	US-09-815-242-4779	Sequence 4779, Ap
43	40	2.2	18846	9	US-09-815-242-8898	Sequence 8898, Ap
44	39.8	2.1	401	9	US-09-815-242-2951	Sequence 2951, Ap
45	39.8	2.1	401	12	US-10-282-122A-5515	Sequence 5515, Ap

ALIGNMENTS

RESULT 1

US-09-564-479-3
; Sequence 3, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: 032931/0230
; CURRENT FILING DATE: 2000-05-03
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1852
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1749)
US-09-564-479-3

Query Match 100.0%; Score 1852; DB 10; Length 1852;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCTCGGAGATAAGCTGGGTATTCCTTCTAGTAACAGCTCGTCTTCTACTACGAGA 60

Db 1 ATGAGCTCGGAGATAAGCTGGGTATTCCTTCTAGTAACAGCTCGTCTTCTACTACGAGA 60

Qy 61 TCTGCAGCGTGGACTCAACGACGCGACCTACGCTCTCTCCACCCACGTTTGTAT 120

[illegible]

RESULT 2
US-09-564-479-1
Sequence 1, Application US/09564479
Publication No. US20030095973A1
GENERAL INFORMATION:
APPLICANT: MURDIN, ANDREW D.
APPLICANT: COOMEN, RAYMOND P.
APPLICANT: WANG, JOE
APPLICANT: DUNN, PAMELA
TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 032931/0230
CURRENT APPLICATION NUMBER: US/09/564,479
CURRENT FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/132,270
PRIOR FILING DATE: 1999-05-03
PRIOR APPLICATION NUMBER: 60/141,276
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2156
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
FEATURE:
NAME/KEY: CDS

LOCATION: (101)..(2053)
US-09-564-479-1

Query Match		99.8%; Score 1848.8; DB 10; Length 2156;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 1850; Conservative		0; Mismatches 2; Indels 0; Gaps 0;
QY	1	ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 60
DB	305	ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA 364
QY	61	TCTGCAGAGCTGGACTCAACGACAGCAGCGACCGACCTACCGCTCCTCCACCCACGTTTGAT 120
DB	365	TCTGCAGAGCTGGACTCAACGACAGCAGCGACCGACCTACCGCTCCTCCACCCACGTTTGAT 424
QY	121	GATTATAAGACTCAAGCGCAACAGCTTACGATFACCTATCTTTACCTCAACATCACTAGCT 180
DB	425	GATTATAAGACTCAAGCGCAACAGCTTACGATFACCTATCTTTACCTCAACATCACTAGCT 484
QY	181	GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAAATATAAGGATACAGCG 240
DB	485	GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAAATATAAGGATACAGCG 544
QY	241	GCTACTGATGAGGAAACCCGCAATCGCTGCGAGTGGGAACTAAGAAATGCCGATGCACTT 300
DB	545	GCTACTGATGAGGAAACCCGCAATCGCTGCGAGTGGGAACTAAGAAATGCCGATGCACTT 604
QY	301	AAAGTTGGCGCGCAAAATTCAGAAATGAGCGAAATATGCTTCGGATAACCAAGCACTTCT 360
DB	605	AAAGTTGGCGCGCAAAATTCAGAAATGAGCGAAATATGCTTCGGATAACCAAGCACTTCT 664
QY	361	GACTCTTTAGTAAACTGACTTCTTCCGACCTCTTACAGGCTGCTTCTTCCAAATCTGTA 420
DB	665	GACTCTTTAGTAAACTGACTTCTTCCGACCTCTTACAGGCTGCTTCTTCCAAATCTGTA 724
QY	421	GCAACAATAAACAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 480
DB	725	GCAACAATAAACAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 784
QY	481	GGGAAACCCCTGCAATTCCTCAATCTTGTAGTGTACAGACAGATGCTACACGCAACAG 540
DB	785	GGGAAACCCCTGCAATTCCTCAATCTTGTAGTGTACAGACAGATGCTACACGCAACAG 844
QY	541	ATAGAGAAAGTGGAAATCGGATTAGGATGATATTTTGGAGCAGACAGACCTAGTGA 600
DB	845	ATAGAGAAAGTGGAAATCGGATTAGGATGATATTTTGGAGCAGACAGACCTAGTGA 904
QY	601	GCTGTAGAAATGCTAAATCTAATAACAGTAAAGCAACATAGATTCAGCTAAAGCAGCA 660
DB	905	GCTGTAGAAATGCTAAATCTAATAACAGTAAAGCAACATAGATTCAGCTAAAGCAGCA 964
QY	661	ATCGTACTGCTTAAGACACAATAGCTGAGCTCAGAAAAGTCCCGGACTCTCCAAT 720
DB	965	ATCGTACTGCTTAAGACACAATAGCTGAGCTCAGAAAAGTCCCGGACTCTCCAAT 1024
QY	721	CTTCAAGAGCGGAAACAAATGTAATACAGCTGAGAAAGATCTTAAATAATCAAACT 780
DB	1025	CTTCAAGAGCGGAAACAAATGTAATACAGCTGAGAAAGATCTTAAATAATCAAACT 1084
QY	781	GCAGATGCTCTGATGTTTCCAAATCCAGGAACACAGTTGGAGGCTCCAAAGCAACAGGA 840
DB	1085	GCAGATGCTCTGATGTTTCCAAATCCAGGAACACAGTTGGAGGCTCCAAAGCAACAGGA 1144
QY	841	AGTAGTATGCTGATGTTTCCAAATCCAGGAACACAGTTGGAGGCTCCAAAGCAACAGGA 900
DB	1145	AGTAGTATGCTGATGTTTCCAAATCCAGGAACACAGTTGGAGGCTCCAAAGCAACAGGA 1204
QY	901	TCCATTTGATCTCGGTTCTGATGATGATTCATGTTCAATACGAAAAATCTCTGAT 960
DB	1205	TCCATTTGATCTCGGTTCTGATGATGATTCATGTTCAATACGAAAAATCTCTGAT 1264
QY	961	TCTCAAGCTGCCAACAGGAGCTCCAGCAACAGCTAGACAGCGAAAGCCGCTGGAGAT 1020

RESULT 3
US-10-289-762-1/c
; Sequence 1, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-995
; CURRENT APPLICATION NUMBER: US/10/289/762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1

DB	1265	TCTCAAGCTGCCAACACAGGAGCTCGACACACAAAGCTAGACGACGAAAGCCGCTGGAGAT 1324
QY	1021	GACAGTGTCTGCTGACGCGTGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
DB	1325	GACAGTGTCTGCTGACGCGTGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1384
QY	1081	GCTGGGCAACACAGGCGATCTCAATGCTTTAGACACAGATCGCTTCTGCTGCTGTG 1140
DB	1385	GCTGGGCAACACAGGCGATCTCAATGCTTTAGACACAGATCGCTTCTGCTGCTGTG 1444
QY	1141	ACCGCAGAGAGTCTCTCCGCTGCAGCAAGTTCATATAGGTCATCTGTAAAAAGCTTTAC 1200
DB	1445	ACCGCAGAGAGTCTCTCCGCTGCAGCAAGTTCATATAGGTCATCTGTAAAAAGCTTTAC 1504
QY	1201	AAGACCTCAAAATCTACAGGTTCTGATTATATAAACACACAGATATCAGAGGTTATGATCT 1260
DB	1505	AAGACCTCAAAATCTACAGGTTCTGATTATATAAACACACAGATATCAGAGGTTATGATCT 1564
QY	1261	TACAAATCAATCAATGATGCTGATGCGGCAAGAAATGATGCGACTCGTATGATGATA 1320
DB	1565	TACAAATCAATCAATGATGCTGATGCGGCAAGAAATGATGCGACTCGTATGATGATA 1624
QY	1321	AACATGTAAGTACCCCGCTCTCAGACGATCCGTTCTAGACGACGACGACGACGACG 1380
DB	1625	AACATGTAAGTACCCCGCTCTCAGACGATCCGTTCTAGACGACGACGACGACGACG 1684
QY	1381	GGACGAGAAAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGACGAACTCTT 1440
DB	1685	GGACGAGAAAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGACGAACTCTT 1744
QY	1441	GGAGATGCTATAGTCAAGTTTGGGCACTACAATCTGTAATGCGAGATCATCCAGTCAAT 1500
DB	1745	GGAGATGCTATAGTCAAGTTTGGGCACTACAATCTGTAATGCGAGATCATCCAGTCAAT 1804
QY	1501	CCTCAAGCGAAATATAGGAGATCAGACAAAGCTTACATCGGAGTCAACAAAGCTCCA 1560
DB	1805	CCTCAAGCGAAATATAGGAGATCAGACAAAGCTTACATCGGAGTCAACAAAGCTCCA 1864
QY	1561	CAGTTTGGCTATCTCTTATGTCGCACTTTCTAATGACTCTACACAGAACTTCATAGCTAAA 1620
DB	1865	CAGTTTGGCTATCTCTTATGTCGCACTTTCTAATGACTCTACACAGAACTTCATAGCTAAA 1924
QY	1621	TTAGAAAGTTTGTGCTGGAAGGATCTAGGACAGAGCTGAAATAAAGCACTTTCCCTTT 1680
DB	1925	TTAGAAAGTTTGTGCTGGAAGGATCTAGGACAGAGCTGAAATAAAGCACTTTCCCTTT 1984
QY	1681	GAAACGAACTCTCTGTTTATTTTATTCAGCAGTGTGCTGCTCAATATCGGCTCTCTATATCTGCT 1740
DB	1985	GAAACGAACTCTCTGTTTATTTTATTCAGCAGTGTGCTGCTCAATATCGGCTCTCTATATCTGCT 2044
QY	1741	TATCTTCAATTAACACACCTTAAGTGTTCGTTTGGAGAGATTTATTTGCTTTGCTAAGG 1800
DB	2045	TATCTTCAATTAACACACCTTAAGTGTTCGTTTGGAGAGATTTATTTGCTTTGCTAAGG 2104
QY	1801	CCTTTGTTGAGGCTTACCAACACACACAGATCTTCAATAAATAAAGA 1852
DB	2105	CCTTTGTTGAGGCTTACCAACACACACAGATCTTCAATAAATAAAGA 2156

[Handwritten signature]

```
,
,
, LENGTH: 1230025
, TYPE: DNA
, ORGANISM: Chlamydia pneumoniae
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (1)..(15000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (15001)..(30000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (30001)..(45000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (45001)..(60000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (60001)..(75000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (75001)..(90000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (90001)..(105000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (105001)..(120000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (120001)..(135000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (135001)..(150000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (150001)..(165000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (165001)..(180000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (180001)..(195000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (195001)..(210000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (210001)..(225000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (225001)..(240000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (240001)..(255000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (255001)..(270000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (270001)..(285000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (285001)..(300000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (300001)..(315000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (315001)..(330000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (330001)..(345000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (345001)..(360000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (360001)..(375000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (375001)..(390000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (390001)..(405000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (405001)..(420000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (420001)..(435000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (435001)..(450000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (450001)..(465000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (465001)..(480000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (480001)..(495000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (495001)..(510000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (510001)..(525000)
, OTHER INFORMATION: n=a or c or g or t
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (525001)..(540000)
```

/	/	OTHER INFORMATION: n=a or c or g or t	Db	828034	GACTCTTTAGGTAAAGTGAATCTCTTCTTCCAGCTCTTCTTACAGGCTGCTCTTCTTCCAAATCTGTA	827975
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	421	GCAAA CAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	480
/	/	LOCATION: (540001)..(550000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827974	GCAAA CAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	827915
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	481	GGGAAAA CGCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACACGACACAG	540
/	/	LOCATION: (555001)..(570000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827914	GGGAAAA CGCCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACACGACACAG	827855
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	541	ATAGAGAAAGATGGAAATGGCAATTAGGGATGATATTTTGACGACAGAACTCTAGTGA	600
/	/	LOCATION: (570001)..(585000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827854	ATAGAGAAAGATGGAAATGGCAATTAGGGATGATATTTTGACGACAGAACTCTAGTGA	827795
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	601	GCTGTAGAAATGCTTAAATCTAATAACAGATTAAGCAACATAGATTCAGCTTAAACAGCA	660
/	/	LOCATION: (585001)..(600000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827794	GCTGTAGAAATGCTTAAATCTAATAACAGATTAAGCAACATAGATTCAGCTTAAACAGCA	827735
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	661	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT	720
/	/	LOCATION: (600001)..(615000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827734	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAATT	827675
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	721	CTTCAAGAAAGCGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAATAATCAAACT	780
/	/	LOCATION: (615001)..(630000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827674	CTTCAAGAAAGCGAAACAAATGGTAATACAGGCTGAGAAAGATCTTAAATAATCAAACT	827615
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	781	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACAGGA	840
/	/	LOCATION: (630001)..(645000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827614	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAAAGCAACAGGA	827555
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	841	AGTAGTATTGGTAGTATTTCGTTGTTTCCATGCTGTTAGATGCTGAAATAGACACCGCT	900
/	/	LOCATION: (645001)..(660000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827554	AGTAGTATTGGTAGTATTTCGTTGTTTCCATGCTGTTAGATGCTGAAATAGACACCGCT	827495
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	901	TCCATTTTTCATGCTCTGGGTTTCGTCAGATGATTCACATGTTCAATACGAAATCTCTGAT	960
/	/	LOCATION: (660001)..(675000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827494	TCCATTTTTCATGCTCTGGGTTTCGTCAGATGATTCACATGTTCAATACGAAATCTCTGAT	827435
/	/	FEATURE:				
/	/	NAME/KEY: misc.feature	QY	961	TCTCAAGCTGCCCAACAGAGCTCCAGCAACAAGCTAGACAGCGAAAGCCGCTCGAGAT	1020
/	/	LOCATION: (675001)..(690000)				
/	/	OTHER INFORMATION: n=a or c or g or t	Db	827434	TCTCAAGCTGCCCAACAGAGCTCCAGCAACAAGCTAGACAGCGAAAGCCGCTCGAGAT	827375
/	/	FEATURE:				
/	/	Query Match				
/	/	Best Local Similarity		99.1%; Score 1836.2; DB 15; Length 1230025;		
/	/	Matches 1849; Conservative		0; Mismatches 3; Indels 1; Gaps 1;		
QY	1	ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	60			
Db	828394	ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	828335			
QY	61	TCTGCAGAGCTGGACTCAACGACAGCGACCGACCTACGCTCCCTCCACCCAGTTTGT	120			
Db	828334	TCTGCAGAGCTGGACTCAACGACAGCGACCGACCTACGCTCCCTCCACCCAGTTTGT	828275			
QY	121	GATTATAAGCTCAAGCGCAAAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT	180			
Db	828274	GATTATAAGCTCAAGCGCAAAACAGCTTACGATCTATCTTTACCTCAACATCACTAGCT	828215			
QY	181	GACATACAGCTGCTTTTGTGAGCTCCAGGATGCTGCTCACTAATATAAGGATACAGCG	240			
Db	828214	GACATACAGCTGCTTTTGTGAGCTCCAGGATGCTGCTCACTAATATAAGGATACAGCG	828155			
QY	241	GCTACTGATGAGGAACCCCAATCGCTGGGAGTGGGAACTAAGAAATGCCATGCAATT	300			
Db	828154	GCTACTGATGAGGAACCCCAATCGCTGGGAGTGGGAACTAAGAAATGCCATGCAATT	828095			
QY	301	AAAGTTGGCGCGCAAAATTACAGAAATAGCGAAATATGCTTCGATAACCAACGATTCTT	360			
Db	828094	AAAGTTGGCGCGCAAAATTACAGAAATAGCGAAATATGCTTCGATAACCAACGATTCTT	828035			
QY	361	GACTCTTTAGGTAAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCAAATCTGTA	420			

```
QY 1500 TCCTCAAGCGGAATATAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAGCCTCC 1559
Db 826894 TCCTCAAGCGGAATATAGGAGATCAGACAAAAGCTTACATCGGAGTGACAAAGCCTCC 826895
QY 1560 ACAGTTGGCTATCTTATGTGCAACCTTCTAATGACTCTACACAGAAGTTCTATAGCTAA 1619
Db 826894 ACAGTTGGCTATCTTATGTGCAACCTTCTAATGACTCTACACAGAAGTTCTATAGCTAA 826775
QY 1620 ATTAGAAAGTTTGTGTGCTGAAGGATCTAGACACAGCTGAATATAAAGCACTTTCTCT 1679
Db 826774 ATTAGAAAGTTTGTGTGCTGAAGGATCTAGACACAGCTGAATATAAAGCACTTTCTCT 826715
QY 1680 TGAACGACACCTCTGTTTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATCTGG 1739
Db 826714 TGAACGACACCTCTGTTTATTCAGCAGGTGCTGCTCAATATCGGCTCTCTATATCTGG 826655
QY 1740 TTATCTCAATAACAAACACCTTAAGTGTTCGTTTGGAGAGATTATATGTGCTTTGGTAAG 1799
Db 826654 TTATCTCAATAACAAACACCTTAAGTGTTCGTTTGGAGAGATTATATGTGCTTTGGTAAG 826595
QY 1800 GCCTTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 1852
Db 826594 GCCTTTGTGAGGCTTACCAACACACTAGAACGATCTTCAATAAATAAAGA 826542

RESULT 4
US-09-841-132-385
; Sequence 385, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yashir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 385
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-132-385

Query Match 94.4%; Score 1748.8; DB 9; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAAACAGCTGCTCTTCTACTAGCAGA 60
Db 205 ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAAACAGCTGCTCTTCTACTAGCAGA 264
QY 61 TCTGACAGCTGGACTCAACGACGACGACGACCTAGCTCTCTCCACCCAGCTTTGAT 120
Db 265 TCTGACAGCTGGACTCAACGACGACGACGACCTAGCTCTCTCCACCCAGCTTTGAT 324
QY 121 GATTATAAGACTCAACGCAACACAGCTTACGATATCTTCTTACCTCAACATCACTAGCT 180
Db 325 GATTATAAGACTCAACGCAACACAGCTTACGATATCTTCTTACCTCAACATCACTAGCT 384
QY 181 GACATACAGGCTGCTTTGGTGGCTCCAGAGTGTGCTCACTATATAAAGATACAGCG 240
Db 385 GACATACAGGCTGCTTTGGTGGCTCCAGAGTGTGCTCACTATATAAAGATACAGCG 444
QY 241 GCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAACTAAGATGCCAGTGCAGTT 300
Db 445 GCTACTGATGAGGAACCGCAATCGCTGGGAGTGGGAACTAAGATGCCAGTGCAGTT 504
QY 301 AAAAGTTGGCGGCAAAATACAGAAATTAGCGAAATATGCTTCGGATAACCAAGCGGATCTT 360
```

```
Db 505 AAAGTTGGCGGCAAAATACAGAAATTAGCGAAATATGCTTCGGANATACCAAGCGATCTTT 564
QY 361 GACTCTTTAGTAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTTCCAATCTGTA 420
Db 565 GACTCTTTAGTAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCTTCCAATCTGTA 624
QY 421 GCAAACTAATCAAAAGCAGCTGAGCTTCTTAAAGAGATGAAGATAAACCCAGTAGTCCCA 480
Db 625 GCAAACTAATCAAAAGCAGCTGAGCTTCTTAAAGAGATGAAGATAAACCCAGTAGTCCCA 684
QY 481 GGGAAACCGCTCAATGCTCAATCTTAAAGAGATGAAGATGAAGATGAAGATGAAGATGAAG 540
Db 685 GGGAAACCGCTCAATGCTCAATCTTAAAGAGATGAAGATGAAGATGAAGATGAAGATGAAG 744
QY 541 ATAGAGAAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 600
Db 745 ATAGAGAAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 804
QY 601 GCTGTAGAAATGCTAAATCTTAATACAGTATGAAGATGAAGATGAAGATGAAGATGAAGATGA 660
Db 805 GCTGTAGAAATGCTAAATCTTAATACAGTATGAAGATGAAGATGAAGATGAAGATGAAGATGA 864
QY 661 ATGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTTCCAATT 720
Db 865 ATGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGGACTCTTCCAATT 924
QY 721 CTTCAAGAACGGGAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 780
Db 925 CTTCAAGAACGGGAACAAATGGTAATACAGGCTGAGAAAGATCTTAAAAATATCAAACT 984
QY 781 GCAGATGCTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACAAGGA 840
Db 985 GCAGATGCTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACAAGGA 1044
QY 841 AGTAGTATTGCTAGTATTCTGTTTCCATGCTGTAGATGATGCTGAAAAATGACACGCT 900
Db 1045 AGTAGTATTGCTAGTATTCTGTTTCCATGCTGTAGATGATGCTGAAAAATGACACGCT 1104
QY 901 TCCATTTTGTATGCTGGTTTGGTTCAGATGATTCATGTTTCAATACGGAATATCTTGAT 960
Db 1105 TCCATTTTGTATGCTGGTTTGGTTCAGATGATTCATGTTTCAATACGGAATATCTTGAT 1164
QY 961 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGTACAGTACAGCAAGCAAGCGCTGGAGAT 1020
Db 1165 TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGTACAGTACAGCAAGCAAGCGCTGGAGAT 1224
QY 1021 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
Db 1225 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1284
QY 1081 GCTGGGCAACAAAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
Db 1285 GCTGGGCAACAAAGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1344
QY 1141 AGCGCAGAGTTCTCTCCGCTGCGACAAAGTTCTATAGGCTCATCTGTAAAAACAGCTTTAC 1200
Db 1345 AGCGCAGAGTTCTCTCCGCTGCGACAAAGTTCTATAGGCTCATCTGTAAAAACAGCTTTAC 1404
QY 1201 AAGACCTCAAAATCTACAGTTCTGATTTATAAACAACAGATATCAGCAGTTATGATGCT 1260
Db 1405 AAGACCTCAAAATCTACAGTTCTGATTTATAAACAACAGATATCAGCAGTTATGATGCT 1464
QY 1261 TACAAATTCGATCAATGCTATGCTAGCGACGAAATGATCGGACTCGTGATGTGATA 1320
Db 1465 TACAAATTCGATCAATGCTATGCTAGCGACGAAATGATCGGACTCGTGATGTGATA 1524
QY 1321 AACAAATTAAGTACCCCGCTCTCAGCAGATCGCTTCTAGACGACGAAACAGAGCTCGA 1380
Db 1525 AACAAATTAAGTACCCCGCTCTCAGCAGATCGCTTCTAGACGACGAAACAGAGCTCGA 1584
QY 1381 GGACCCAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1440
Db 1585 GGACCCAGAAAAACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1644
```

QY 1441 GGAGATGCTATAGTCAAGTTTGGGACTACAACTCTGTATGAGATCACTCAGTCGAAT 1500
Db 1645 GGAGATGCTATAGTCAAGTTTGGGACTACAACTCTGTATGAGATCACTCAGTCGAAT 1704
QY 1501 CCTCAAGCAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCCA 1560
Db 1705 CCTCAAGCAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCCA 1764
QY 1561 CAGTTTGGCTATCCTTATGTAAGCACTTCTTAATGACTCTACACAGAAAGTTTCAATAGCTAAA 1620
Db 1765 CAGTTTGGCTATCCTTATGTAAGCACTTCTTAATGACTCTACACAGAAAGTTTCAATAGCTAAA 1824
QY 1621 TTAGAAAGTTTGTGCTGAAGATCTAGACAGCAGCTGAAATATAAGCACTTCTCTTT 1680
Db 1825 TTAGAAAGTTTGTGCTGAAGATCTAGACAGCAGCTGAAATATAAGCACTTCTCTTT 1884
QY 1681 GAAACGAACCTCTTGTATTATCAGCAGGTGCTGCTCAATATCGGCTCTCTATATCTGGT 1740
Db 1885 GAAACGAACCTCTTGTATTATCAGCAGGTGCTGCTCAATATCGGCTCTCTATATCTGGT 1944
QY 1741 TATCTCCAAATAA 1752
Db 1945 TATCTCCAAATAA 1956

RESULT 5
US-09-841-260-63
; Sequence 63, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Brika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-09-841-260-63

Query Match 94.4%; Score 1748.8; DB 10; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGCGAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
Db 205 ATGAGTCTGCGAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264
QY 61 TTGCGAGCGTGGACTCAACGACGAGCGACCGCACTTACGCTCTCTCCACCCAGCTTTGAT 120
Db 265 TTGCGAGCGTGGACTCAACGACGAGCGACCGCACTTACGCTCTCTCCACCCAGCTTTGAT 324
QY 121 GATTATAAGACTCAAGCGGAAACAGCTTACGATATCTTTACCTTCAACATCAGCTAGCT 180
Db 325 GATTATAAGACTCAAGCGGAAACAGCTTACGATATCTTTACCTTCAACATCAGCTAGCT 384
QY 181 GACATACAGGCTCTTTTGTGTAGCTTCCAGGATGCTGCTCACTAATAATAAGGATACAGCG 240
Db 385 GACATACAGGCTCTTTTGTGTAGCTTCCAGGATGCTGCTCACTAATAATAAGGATACAGCG 444
QY 241 GCTACTGATGAGAAACCCCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGCGATT 300
Db 445 GCTACTGATGAGAAACCCCAATCGCTGGGAGTGGGAAACTAAGAAATGCCGATGCGATT 504
QY 301 AAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCAATCTTT 360
Db 505 AAAGTTGGCGCGCAATATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCAATCTTT 564

QY 1441 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAAATGCGATCACTCAGTCGAAT 1500
Db 1645 GGAGATGCTATAGTCAAGTTTCGGCACTACAATCTGTAAATGCGATCACTCAGTCGAAT 1704
QY 1501 CCTCAGCGAATAATGAGGATCAGACAAAGTTTACATCGGAGTGCAGAAAGCTTCCA 1560
Db 1705 CCTCAGCGAATAATGAGGATCAGACAAAGTTTACATCGGAGTGCAGAAAGCTTCCA 1764
QY 1561 CAGTTTGGCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTTCATAGCTAAA 1620
Db 1765 CAGTTTGGCTATCCTTATGTGCAACTTTCTAATGACTCTACACAGAAAGTTTCATAGCTAAA 1824
QY 1621 TTGAAAGTTTGTCTGCTGAGGATCTAGGACAGCAGCTGAAATTAAGAGCTTTTCTTCTTT 1680
Db 1825 TTGAAAGTTTGTCTGCTGAGGATCTAGGACAGCAGCTGAAATTAAGAGCTTTTCTTCTTT 1884
QY 1681 GAAACGAACTCTTGTCTTATTCAGCAGCTGCTGCTCAATATCGGCTCTCTATATCTGGT 1740
Db 1885 GAAACGAACTCTTGTCTTATTCAGCAGCTGCTGCTCAATATCGGCTCTCTATATCTGGT 1944
QY 1741 TATCTCCAATAA 1752
Db 1945 TATCTCCAATAA 1956

RESULT 6

US-10-007-693-63
; Sequence 63, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 63
; LENGTH: 1956
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-63

Query Match 94.4%; Score 1748.8; DB 13; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAAGCTCTCTTCTACTAGCAGA 60
Db 205 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAAAGCTCTCTTCTACTAGCAGA 264
QY 61 TCTGAGAGCTGGAATCAACGACGAGCGACCGACCTAGCCCTCCACCCACGTTTGTAT 120
Db 265 TCTGAGAGCTGGAATCAACGACGAGCGACCGACCTAGCCCTCCACCCACGTTTGTAT 324
QY 121 GATTATAGACTCAAGCGCAACACAGCTTACGATACATCTTCTTACCTCAACATCACTAGT 180
Db 325 GATTATAGACTCAAGCGCAACACAGCTTACGATACATCTTCTTACCTCAACATCACTAGT 384
QY 181 GACATACAGGCTGCTTGTGTAGCTTCCAGGATGCTGCTCACTAATATAAGGATACAGCG 240
Db 385 GACATACAGGCTGCTTGTGTAGCTTCCAGGATGCTGCTCACTAATATAAGGATACAGCG 444
QY 241 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCGGATCGAGTT 300
Db 445 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCGGATCGAGTT 504
QY 301 AAGTTGGCGCGCAATATACAGATTTAGCGAAATATGCTTCGGATTAACCAAGCGATTTCTT 360
Db 505 AAGTTGGCGCGCAATATACAGATTTAGCGAAATATGCTTCGGATTAACCAAGCGATTTCTT 564
QY 361 GACTCTTTAGGTAAGTGAATCTCTTCCAGCTTACAGGCTGCTTCTTCCAAATCTGTA 420

Db 565 GACTCTTTAGGTAAGTGAATCTCTTCCAGCTTACAGGCTGCTTCTTCCAAATCTGTA 624
QY 421 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATCAACCCAGTAGTCCCA 480
Db 625 GCAAAACAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATCAACCCAGTAGTCCCA 684
QY 481 GGGAAAAACGCTCGCAATGCTCAATCTTATAGTTGATCAGACAGATGCTACAGGACACAG 540
Db 685 GGGAAAAACGCTCGCAATGCTCAATCTTATAGTTGATCAGACAGATGCTACAGGACACAG 744
QY 541 ATAGAGAAAGATGGAATGCGATGAGGATGCATATTTTCGAGGACAGACGCTAGTGGGA 600
Db 745 ATAGAGAAAGATGGAATGCGATGAGGATGCATATTTTCGAGGACAGACGCTAGTGGGA 804
QY 601 GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGCAGCA 660
Db 805 GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTTCAGCTAAAGCAGCA 864
QY 661 ATGCTACTGCTAAGACACAAAATAGCTGAAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 720
Db 865 ATGCTACTGCTAAGACACAAAATAGCTGAAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 924
QY 721 CTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAGAAAGATCTTAAATAATATCAAACT 780
Db 925 CTTCAAGAAGCGGAACAAATGGTAATACAGGCTGAGAAAGATCTTAAATAATATCAAACT 984
QY 781 GCAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGCTCCAAAGCAACAGGA 840
Db 985 GCAGATGCTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGCTCCAAAGCAACAGGA 1044
QY 841 AGTAGTATGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 900
Db 1045 AGTAGTATGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT 1104
QY 901 TCCATTTTGTATGCTGCTGGTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCTGAT 960
Db 1105 TCCATTTTGTATGCTGCTGGTTTCGTCAGATGATTCACATGTTCAATACGGAATAATCTGAT 1164
QY 961 TCTCAAGCTGCCAACAGGAGCTGCGACGCAAGCTAGAGCAGGAAAGCGCTGGAGAT 1020
Db 1165 TCTCAAGCTGCCAACAGGAGCTGCGACGCAAGCTAGAGCAGGAAAGCGCTGGAGAT 1224
QY 1021 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1080
Db 1225 GACAGTCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA 1284
QY 1081 GCTGGCAACAAACAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1140
Db 1285 GCTGGCAACAAACAGGCGTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG 1344
QY 1141 AGCAGAGGTTCTCCGCTGCGAGAAAGTTCTATAGGTCATCTGTAAACAGCTTTAC 1200
Db 1345 AGCAGAGGTTCTCCGCTGCGAGAAAGTTCTATAGGTCATCTGTAAACAGCTTTAC 1404
QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTATGATGCT 1260
Db 1405 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTATGATGCT 1464
QY 1261 TACAAATCCATCAATGATGCTTGGTAGGCGCAAAATGATGCGACTCGTGTGTGATA 1320
Db 1465 TACAAATCCATCAATGATGCTTGGTAGGCGCAAAATGATGCGACTCGTGTGTGATA 1524
QY 1321 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAAGAAAGAGCTCGA 1380
Db 1525 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAAGAAAGAGCTCGA 1584
QY 1381 GGACAGAAAAAAGATCAGACCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1440
Db 1585 GGACAGAAAAAAGATCAGACCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTT 1644
QY 1441 GGAGATGCTATAGTCAAGTTTCGGCACTCAAACTCTGTAAATGCGATCACTCAGTCGAAT 1500

Db 1645 GGAGATGCTATAGTCAAGTTTCGGCACTACATCTGTAATGCAGATCATCCAGTCAAT 1704
 QY 1501 CCTCAAGCGAATATAGAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCTTCCA 1560
 Db 1705 CCTCAAGCGAATATAGAGGAGATCAGACAAAGCTTACATCGGCGAGTGACAAAGCTTCCA 1764
 QY 1561 CAGTTTGGCTATCCTTATGTCGAATCTTCTAATGACTCTACACAGAGAGTTTCATAGCTAAA 1620
 Db 1765 CAGTTTGGCTATCCTTATGTCGAATCTTCTAATGACTCTACACAGAGAGTTTCATAGCTAAA 1824
 QY 1621 TTAGAAAGTTTGGTTGCTGAAGGATCTAGGACAGCAGCTGAAATPAAAGACACTTTCCTTT 1680
 Db 1825 TTAGAAAGTTTGGTTGCTGAAGGATCTAGGACAGCAGCTGAAATPAAAGACACTTTCCTTT 1884
 QY 1681 GAAACGAACTCCCTTGTATTATCAGCAGGCTGCTGGTGAATATCGGCTCTCTATATTCGGT 1740
 Db 1885 GAAACGAACTCCCTTGTATTATCAGCAGGCTGCTGGTGAATATCGGCTCTCTATATTCGGT 1844
 QY 1741 TATCTCCAATAA 1752
 Db 1945 TATCTCCAATAA 1956

RESULT 7

US-10-312-273-14
 ; Sequence 14, Application US/10312273
 ; Publication No. US20040005667A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON Spa
 ; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
 ; FILE REFERENCE: P025035W0
 ; CURRENT APPLICATION NUMBER: US/10/312-273
 ; CURRENT FILING DATE: 2002-12-20

; PRIOR APPLICATION NUMBER: 061563-4
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 0017047.2
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 0017983.8
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: 0019368.0
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: 0020440.4
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: 0022583.9
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 0027549.5
 ; PRIOR FILING DATE: 2000-11-10
 ; PRIOR APPLICATION NUMBER: 0031706.5
 ; PRIOR FILING DATE: 2000-12-22
 ; NUMBER OF SEQ ID NOS: 664
 ; SOFTWARE: SeqWin99, version 1.02
 ; SEQ ID NO 14

; LENGTH: 1956
 ; TYPE: DNA

; ORGANISM: Chlamydia pneumoniae

US-10-312-273-14

Query Match 94.4%; Score 1748.8; DB 15; Length 1956;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
 Db 205 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264
 QY 61 TCTGCAGAGCTGGACTCAACGACAGCGACCGACCTTACGCTCTCCACCCACGTTTGTAT 120
 Db 265 TCTGCAGAGCTGGACTCAACGACAGCGACCGACCTTACGCTCTCCACCCACGTTTGTAT 324
 QY 121 GATTATAAGACTCAACGCGAAGAGCTTAGGATATCTTTACCTTCAACATCACTAGCT 180
 Db 325 GATTATAAGACTCAACGCGAAGAGCTTAGGATATCTTTACCTTCAACATCACTAGCT 384

QY 181 GACATACAGGCTGCTTTGGTGGAGCCCTCCAGGATGCTGTCTCACTAATATAAAGGATACAGCG 240
 Db 385 GACATACAGGCTGCTTTGGTGGAGCCCTCCAGGATGCTGTCTCACTAATATAAAGGATACAGCG 444
 QY 241 GCTACTGATGAGGAAACCCCAATCGCTGGGAGTGGAACTAAGAAATGCCGATGAGTT 300
 Db 445 GCTACTGATGAGGAAACCCCAATCGCTGGGAGTGGAACTAAGAAATGCCGATGAGTT 504
 QY 301 AAAGTTGGCGCGCAAAATTACAGAAATTAGGCAATATGCTTCGGATAAACCAAGCGATTCTTT 360
 Db 505 AAAGTTGGCGCGCAAAATTACAGAAATTAGGCAATATGCTTCGGATAAACCAAGCGATTCTTT 564
 QY 361 GACTCTTTAGGTAACTGACTTCTTCTCGACCTCTTACAGGCTCTCTTCTCAATCTGTA 420
 Db 565 GACTCTTTAGGTAACTGACTTCTTCTCGACCTCTTACAGGCTCTCTTCTCAATCTGTA 624
 QY 421 GCAAAACAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGTCCCA 480
 Db 625 GCAAAACAATAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAAACCCAGTAGTCCCA 684
 QY 481 GGGAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCAGACAG 540
 Db 685 GGGAAAACGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTTACAGCAGACAG 744
 QY 541 ATAGCAAAAGATGGAATCGGATTAGGGATGCAATATTTTGCAGGACAGAACGCTAGTGA 600
 Db 745 ATAGCAAAAGATGGAATCGGATTAGGGATGCAATATTTTGCAGGACAGAACGCTAGTGA 804
 QY 601 GCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATAGATTTCAGTAAAGCAGCA 660
 Db 805 GCTGTAGAAAATGCTAAATCTTAATACAGTATAAGCAACATAGATTTCAGTAAAGCAGCA 864
 QY 661 ATCGCTACTGCTAAAGACACAAATAGCTGAAGCTCAGAAAAGTTCGCCGACTCTCCAAAT 720
 Db 865 ATCGCTACTGCTAAAGACACAAATAGCTGAAGCTCAGAAAAGTTCGCCGACTCTCCAAAT 924
 QY 721 CTTCAAGAGCGGAAACAAATGCTTAACAGGCTGAGAAAGATCTTAAATAATATCAAACT 780
 Db 925 CTTCAAGAGCGGAAACAAATGCTTAACAGGCTGAGAAAGATCTTAAATAATATCAAACT 984
 QY 781 GCAGATGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGCA 840
 Db 985 GCAGATGTTCTGATGTTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGCA 1044
 QY 841 ACTAGTATGTTAGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 900
 Db 1045 AGTAGTATGTTAGTATGCTGTTTCCATGCTGTTAGATGATGCTGAAAATGAGACCGCT 1104
 QY 901 TCCATTTTATGCTGCTGGGTTTGGTTCAGATGATTCACATGTTCAATACGGAATCTGAT 960
 Db 1105 TCCATTTTATGCTGCTGGGTTTGGTTCAGATGATTCACATGTTCAATACGGAATCTGAT 1164
 QY 961 TCTCAAGCTGCCCAACAGGAGCTCCGACGACCAAGCTAGACGACGAAAGCCGCTGGAGAT 1020
 Db 1165 TCTCAAGCTGCCCAACAGGAGCTCCGACGACCAAGCTAGACGACGAAAGCCGCTGGAGAT 1224
 QY 1021 GACATGCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGGTAAA 1080
 Db 1225 GACATGCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTCTAGGTAAA 1284
 QY 1081 GCTGGGCAACACAGGCGATCTCAATGCTTTAGACAGATGCTTCTGCTGCTGTTG 1140
 Db 1285 GCTGGGCAACACAGGCGATCTCAATGCTTTAGACAGATGCTTCTGCTGCTGTTG 1344
 QY 1141 AGCGCAGAGTCTTCTCCCGCTGCAGCAAGTCTTAGGGTCACTCTGTAACCAAGCTTTTAC 1200
 Db 1345 AGCGCAGAGTCTTCTCCCGCTGCAGCAAGTCTTAGGGTCACTCTGTAACCAAGCTTTTAC 1404
 QY 1201 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGAGGTTATGATGCT 1260
 Db 1405 AAGACCTCAAAATCTACAGGTTCTGATTAATAAACACAGATATCAGAGGTTATGATGCT 1464
 QY 1261 TACAAATCCATGATGCTGCTTATGGTAGGCGACGAAATGATGCGACTCTGATGTGATA 1320

Db 1465 TACAAATCCATCAATGATGCCCTATGCTAGGACGAAATGATGCGACTCGTATGATA 1524
Qy |||||
Db 1321 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGCTCGA 1380
Qy |||||
Db 1525 AACAAATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAAGAGCTCGA 1584
Qy |||||
Db 1381 GGACAGAAAAAACAGATCAAGCCCTCGCTAGGGTGAATTTCTGGCAATAGCAGAACTCTT 1440
Qy |||||
Db 1585 GGACAGAAAAAACAGATCAAGCCCTCGCTAGGGTGAATTTCTGGCAATAGCAGAACTCTT 1644
Qy |||||
Db 1441 GGAGATGCTATAGTCAAGTTTCGCACTACATCTGTAATGACATCACTCAGTCAAT 1500
Qy |||||
Db 1645 GGAGATGCTATAGTCAAGTTTCGCACTACATCTGTAATGACATCACTCAGTCAAT 1704
Qy |||||
Db 1501 CCTCAAGCGAATAATGAGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTCCA 1560
Qy |||||
Db 1705 CCTCAAGCGAATAATGAGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTCCA 1764
Qy |||||
Db 1561 CAGTTTGGCTATCCCTATGTCGAACTTTCTAATGACTCTACACAGAACTTCACTAGCTAAA 1620
Qy |||||
Db 1765 CAGTTTGGCTATCCCTATGTCGAACTTTCTAATGACTCTACACAGAACTTCACTAGCTAAA 1824
Qy |||||
Db 1621 TTAGAAAGTTTCTTGGCTGAAGATCTAGGACAGCAGCTGAAATAAAGCACTTTCCCTTT 1680
Qy |||||
Db 1825 TTAGAAAGTTTCTTGGCTGAAGATCTAGGACAGCAGCTGAAATAAAGCACTTTCCCTTT 1884
Qy |||||
Db 1681 GAAACGAACCTCTCTTTATTACAGAGTGTGCTCAATATCGGCTCTCTATATCTGGT 1740
Qy |||||
Db 1885 GAAACGAACCTCTCTTTATTACAGAGTGTGCTCAATATCGGCTCTCTATATCTGGT 1944
Qy |||||
Db 1741 TATCTCCAATAA 1752
Qy |||||
Db 1945 TATCTCCAATAA 1956

RESULT 8

US-03-564-479-7
; Sequence 7, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: OMEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, PAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (766)..(2235)
US-09-564-479-7

Query Match 62.7%; Score 1161.8; DB 10; Length 2238;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1174; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 ATGAGTCTGGCAGATAAGCTGGTATTCTTCTAGTAAACAGCTGCTTCTACTAGCAGA 60
Db 970 ATGAGTCTGGCAGATAAGCTGGTATTCTTCTAGTAAACAGCTGCTTCTACTAGCAGA 1029

Qy 61 TCTCAGACGCTGGACTCAACGACAGCAGCAGCAGCTAGCGCTCTCCACCCACCGTTTGAAT 120
Db 1030 TCTCAGACGCTGGACTCAACGACAGCAGCAGCAGCTAGCGCTCTCCACCCACCGTTTGAAT 1089
Qy 121 GATTATTAAGACTCAAGCGCAAAAGAGCTTAAGTACTATCTTTACTCTAAATCACTAGCT 180
Db 1090 GATTATTAAGACTCAAGCGCAAAAGAGCTTAAGTACTATCTTTACTCTCAACATCACTAGCT 1149
Qy 181 GACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTCTCACTAATAATAAGAGATACAGCG 240
Db 1150 GACATACAGGCTGCTTTGGTGGAGCTCCAGGATGCTCTCACTAATAATAAGAGATACAGCG 1209
Qy 241 GCTACTGATAGGAAAAACGCAATCGCTGGAGTGGAAAACTAAGAAATGCGGATGCGAGTT 300
Db 1210 GCTACTGATAGGAAAAACGCAATCGCTGGAGTGGAAAACTAAGAAATGCGGATGCGAGTT 1269
Qy 301 AAGCTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTTCTT 360
Db 1270 AAGCTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATACCAAGCGATTTCTT 1329
Qy 361 GACTCTTTAGTAACTGACTTCTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 420
Db 1330 GACTCTTTAGTAACTGACTTCTCGACCTCTTACAGGCTGCTCTTCTCCAATCTGTA 1389
Qy 421 GCAAACTAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db 1390 GCAAACTAATAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 1449
Qy 481 GGGAAAAACGCTGCAATTTGCTCAATCTTCTAGTCTAGCAGATGCTTACAGCGACACAG 540
Db 1450 GGGAAAAACGCTGCAATTTGCTCAATCTTCTAGTCTAGCAGATGCTTACAGCGACACAG 1509
Qy 541 ATAGAGAAAGATGGAATGCAATTTAGGATGCAATTTTTCAGGACAGAAAGCTAGTGGGA 600
Db 1510 ATAGAGAAAGATGGAATGCAATTTAGGATGCAATTTTTCAGGACAGAAAGCTAGTGGGA 1569
Qy 601 GCTGTAGAAAATGCTAATCTAATAACAGTATAAGCAACATAGATTGAGCTAAAGCAGCA 660
Db 1570 GCTGTAGAAAATGCTAATCTAATAACAGTATAAGCAACATAGATTGAGCTAAAGCAGCA 1629
Qy 661 ATCGCTACTGCTAAGACACAAATAGCTGAAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 720
Db 1630 ATCGCTACTGCTAAGACACAAATAGCTGAAAGCTCAGAAAAAGTTCCCGGACTCTCCAAAT 1689
Qy 721 CTTCAAGACGCGAACAAATGGTAATAACAGCTGAGAAAGATCTTAAAAATATCAAAACCT 780
Db 1690 CTTCAAGACGCGAACAAATGGTAATAACAGCTGAGAAAGATCTTAAAAATATCAAAACCT 1749
Qy 781 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGACACAGGA 840
Db 1750 GCAGATGTTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGACACAGGA 1809
Qy 841 AGTAGTATTGCTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 900
Db 1810 AGTAGTATTGCTAGTATTTCGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 1869
Qy 901 TCCATTTTGTATGCTGGTTTCGTCAGATGATTCATGTTCAATGTCGAAAAATTCCTGAT 960
Db 1870 TCCATTTTGTATGCTGGTTTCGTCAGATGATTCATGTTCAATGTCGAAAAATTCCTGAT 1929
Qy 961 TCTCAAGCTGCCAAACAGGAGCTCGCAGACACAGCTAGAGCAGGAAAGCCGCTGGAGAT 1020
Db 1930 TCTCAAGCTGCCAAACAGGAGCTCGCAGACACAGCTAGAGCAGGAAAGCCGCTGGAGAT 1989
Qy 1021 GACAGTCTCTGACGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTCTAGSTAAA 1080
Db 1990 GACAGTCTCTGACGCGCTGGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTCTAGSTAAA 2049
Qy 1081 GCTGGGGCAACACAGGCGATACATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
Db 2050 GCTGGGGCAACACAGGCGATACATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 2109
Qy 1141 AGCGCAGGAGTTCTCTCCCGCTGCAGCAAGTTCTATAG 1177

DB 2110 AGCGCAGGAG-TACTCCGCTGCAGCAAGTTCTATGG 2145

RESULT 9

US-09-564-479-5
 ; Sequence 5, Application US/09564479
 ; Publication NO. US2003005973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: COMEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; CURRENT FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/1132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR APPLICATION NUMBER: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1456
 ; TYPE: DNA
 ; ORGANISM: Chlamydia pneumoniae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (101)..(1456)
 US-09-564-479-5

```
Query Match      62.1%; Score 1151; DB 10; Length 1456;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA	60
Db	305	ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA	364
QY	61	TCTGCAGACGTGGACTCAACACAGACGACCGCACCTACCGCTCTCCACCCACGTTTGAT	120
Db	365	TCTGCAGACGTGGACTCAACACAGACGACCGCACCTACCGCTCTCCACCCACGTTTGAT	424
QY	121	GATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTTACTCTCAACATCACTAGCT	180
Db	425	GATTATAAGACTCAAGCGCAACAGCTTACGATACTATCTTTTACTCTCAACATCACTAGCT	484
QY	181	GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATAATAAGGATACACGC	240
Db	485	GACATACAGGCTGCTTTGGTGAGCTCCAGGATGCTGTCACTAATAATAAGGATACACGC	544
QY	241	GCTACTGATGAGGAAACCGCAATCGCTCCGAGTGGGAACTAAGAATGCCGATGCAGTT	300
Db	545	GCTACTGATGAGGAAACCGCAATCGCTCCGAGTGGGAACTAAGAATGCCGATGCAGTT	604
QY	301	AAAGTTGGCGGCAAAATTACAGAATTAGCGAAATGCTTCGGGATAACCAAGCGATTCTT	360
Db	605	AAAGTTGGCGGCAAAATTACAGAATTAGCGAAATGCTTCGGGATAACCAAGCGATTCTT	664
QY	361	GACTCTTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA	420
Db	665	GACTCTTTTAGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTCTGTA	724
QY	421	GCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	480
Db	725	GCAACCAATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	784
QY	481	GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGCACACAG	540
Db	785	GGGAAACCGCTGCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGGCACACAG	844

541	QY	ATAGAGAAAGATGGGAAATCGGATTCGGATATGCGATATTTTGCAGGACAGAACGTCGTAGTGA	600
845	Db	ATAGAGAAAGATGGGAAATCGGATTCGGATATTTTGCAGGACAGAACGTCGTAGTGA	904
601	QY	GCTGTAGAAAATGCTAAATCTTAATAACAGTATATAAGCAACATAGATTTCAGCTAAAGCAGCA	660
905	Db	GCTGTAGAAAATGCTAAATCTTAATAACAGTATATAAGCAACATAGATTTCAGCTAAAGCAGCA	964
661	QY	ATCGCTACTGCTAAGACACAAAATAGCTGAAGCTCGAAAAAAGTTCCCGGACTCTCCAATT	720
965	Db	ATCGCTACTGCTAAGACACAAAATAGCTGAAGCTCGAAAAAAGTTCCCGGACTCTCCAATT	1024
721	QY	CTTCAAGAACGGGAACAAATGGTATACAGGCTGGAAGAAGATCTTTAAAAATATCAAAACCT	780
1025	Db	CTTCAAGAACGGGAACAAATGGTATACAGGCTGGAAGAAGATCTTTAAAAATATCAAAACCT	1084
781	QY	GCAGATGGTTCTGATGCTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACAAGGA	840
1085	Db	GCAGATGGTTCTGATGCTTCCAAATCCAGGAACCTACAGTTGGAGGCTCCAGCAACAAGGA	1144
841	QY	AGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTTAGATGATGCTGAAAATAGAGCCGCT	900
1145	Db	AGTAGTATTGGTAGTATTCGTGTTTCCATGCTGTTTAGATGATGCTGAAAATAGAGCCGCT	1204
901	QY	TCCATTTTGATGCTCTGGGTTTCGTACAGATGATTCACATGTTCAATACGGGAAAATCCTGAT	960
1205	Db	TCCATTTTGATGCTCTGGGTTTCGTACAGATGATTCACATGTTCAATACGGGAAAATCCTGAT	1264
961	QY	TCTCAAGCTGCCCAACAGAGAGCTCGCAGCACAAAGCTAGAGCAGCGAAAGCCGCTGGAGAT	1020
1265	Db	TCTCAAGCTGCCCAACAGAGAGCTCGCAGCACAAAGCTAGAGCAGCGAAAGCCGCTGGAGAT	1324
1021	QY	GACAGTCTGCTGCAGCGCTGGCAGATGCTCGAAAAAGCTTTAGAAGCGGCTCTAGGTAAA	1080
1325	Db	GACAGTCTGCTGCAGCGCTGGCAGATGCTCGAAAAAGCTTTAGAAGCGGCTCTAGGTAAA	1384
1081	QY	GCTGGGCAACAACAGGGGCTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTGTG	1140
1385	Db	GCTGGGCAACAACAGGGGCTACTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTGTG	1444
1141	QY	AGCGCAGGAGT 1151	
1445	Db	AGCGCAGGAGT 1455	
RESULT 10			
US-09-841-132-57			
; Sequence 57, Application US/09841132			
; Patent No. US20020061848A1			
; GENERAL INFORMATION:			
; APPLICANT: Bhatia, Ajay			
; APPLICANT: Skeiky, Yasir A.W.			
; APPLICANT: Probst, Peter			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND			
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION			
; FILE REFERENCE: 210121.469C8			
; CURRENT APPLICATION NUMBER: US/09/841,132			
; CURRENT FILING DATE: 2001-04-23			
; NUMBER OF SEQ ID NOS: 599			
; SOFTWARE: Fast-Seq for Windows Version 3.0/4.0			
; SEQ ID NO 57			
; LENGTH: 1537			
; TYPE: DNA			
; ORGANISM: Chlamydia trachomatis			
US-09-841-132-57			

479

QY	923	GT	CAGATGATTCACATGTTCCAAATACGGAAATTCCTGATTCCTCAAGCTGCCCAACAGAGAGC	982
DB	246	GATCTATGATCGAAACAATTTAATGTAAACAATCTCTGCAACAGCTTAAAGAGCTACAAGCTA	305	
QY	983	TCCGAGCACAAAGCTTAGAGACGCGAAAGCCGCTGCAGATGACAGTGCCTGCTGCAGCGCTGG	1042	
DB	306	TGGAGGCTCAGTCA---CTGCGATGTCAAGTCAACTGTTGTTGCTCCGATGGCGAGCTCC	362	
QY	1043	CAGATGCTCAGAAAGCTTTAGAAGCGGCTCTAGTATAAGCTGGGCAACAAC-----	1093	
DB	363	CAGCCGAAATACAGCAATCAAGATGCTCTTGGCCAAGCTTTGAACAACCATCAGCAG	422	
QY	1094	AGGCATACTCAATGCTTTTAGGACAGATCGCTTCCTGCTGCTGTTGTGTGAGCGCAGAGTTC	1153	
DB	423	ATGCTTTAGCTTACAGCTATGGCAAGTGGCTTTTGCACTGCCAAGGTTGGAGGAGGCT	482	
QY	1154	CTCCCGCTGCAGCAAGTCTTATAGGGTCACTGTGTAAAAACAGCTTTTCAAGAGCCTCAAAAT	1213	
DB	483	CCGCAGAAACAGCTGGGCACCTGTCCAGATGAATGTAAACAGCTTTTCAAGACAGCGTTTT	542	
QY	1214	CTACAGGTTCT-----GATTATAAAACACAGATATCAGCAGGTTATGATGCTTTACAAAT	1267	
DB	543	CTTCGACTTCTCCAGCTCTTATGCAGCAGCACITTCOGATGGATATTCCTGCTTACAAAA	602	
QY	1268	CCATCAATGATGCCTATGTTAGGCGACGAAATGATGGACATCTGCTGATGTGATAACAAATG	1327	
DB	603	CATCGAACTCTTTATATTCGAAAGCAGAGCGCGTGCAGTCACTGCTATTAGTCAAACTG	662	
QY	1328	TAAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGAGGACCAAG	1387	
DB	663	CAAAATCCGCGCTTCCAGAGCGTTTCTCGTCTGGCATAGAAGTCAAGGACGCAAGTG	722	
QY	1388	AAAAACAGATCAAGCCCTCGCTAGGGTGATTTCTGGCAATAGCAGAACTCTTTGAGAGATG	1447	
DB	723	CAGATGTAGCCAAAGACGACGAGAAACTATTGTACAGATAGCCAAACGTTAGGTGATG	782	
QY	1448	TCTATAGTCAAGTTTCGGCACTACAACTGTAAATGACAGATCACTCAGTCAATCTCAAG	1507	
DB	783	TATATAGCCGCTTACAGGTTCTGGATTCTTTGATGTTCTAGATTGTGAGCAATCCGCAAG	842	
QY	1508	CGAATAATAGGAGATCAGACAAAGCTTACATCGGCAGTGACAAAGCTCCACAGTTTG	1567	
DB	843	CAAAATCAGABAGATTTATCGAAGAGCTCACGGCATCTATTAGCAAGCTCCACAAATTG	902	
QY	1568	GCTATCCTTATGTGCACTTTCTTAATGACTCTACACAGAGTTCATAGCTAAATTAGAAA	1627	
DB	903	GGTATCCTGTGTTTCAGAAATCTGTGGATAGCTTTCGCAAGGTTTGTCTGCAATTTGAAA	962	
QY	1628	GTTTGTTTGTGAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTCCTTTGAAACGA	1687	
DB	963	GAGAGTTTGTTGATGGGAAACGTAGTCTCGCAGAAATCTCAAGAGAAATGCGTTTAGAAAAC	1022	
QY	1688	ACTCCTTGTTTTATTCAGCAGGTGCTGGTCAATATCGGCTCTCTATATTTCTGTTATCTCC	1747	
DB	1023	AGCCCGCTTTCATTCACAGGTGTGGTAACAATGCTTCTCTATTTCTCTGTTATCTTT	1082	
QY	1748	AATTAAC	1753	
DB	1083	CTTAAAC	1088	

SEQUENCE NO.	SEQUENCE	SEQUENCE NO.	SEQUENCE
1004	CGAAGCCGCTGGAGATGACACATGCTGTCAGCGCGCTGGCAGATGCTCAGAAAGCTTTAG	1064	AAGCGGCTCTAGGTTAAAGCTGGGCAACAAC-----AGGGCATCTCAATGCTTTTAG
1181	CGATGTCAGATCAACTGGTTGGTGGGATGCGAGCTCCGACCGGAATACAGCAATCA	1241	AAGATGCTCTTCGCGAAGCTTTGAAACAACCATCAACAGATGGTTTAGCTACAGCTATGG
1115	GACAGATCGCTTCTGCTGTGTGTGAGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCTTA	1301	GACAAGTGGCTTTTTCAGCTGCCAAGGTTGGAGGAGGCTCCGACAGGAACAGCTGGCACTG

[illegible]

Db 1533 GGTATCCCTGCTGTTTCTCAGAAATCTCGGATAGCTTCGAGAAATTTGGAAA 1592
QY 1628 GTTTGTTTGTGCTGAAGGATCTAGACAGCAGCTGAAATAAAAGCACTTTCCITTTGAAACGA 1687
Db 1593 GAGAGTTTGTGATGGGAACTAGTCTCGCAGAAATCTCAAGAGAAATGGTTTGAAGAAC 1652
QY 1688 ACTCCTGTTTATTCAGCAGAGTGTGCTCAATATCGGCTCTCTATATTCGTTATCTCC 1747
Db 1653 AGCCCGCTTTCATTCAACAGAGTGTGGTAAACAATGCTTCTATCTCTGTTATCTTT 1712
QY 1748 AATAAC 1753
Db 1713 CTTAAC 1718

RESULT 15

US-10-007-693-38
; Sequence 38, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 38
; LENGTH: 1834
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis
US-10-007-693-38

Query Match 9.9%; Score 184; DB 13; Length 1834;

Best Local Similarity 52.8%; Pred No. 1.6e-41;
Matches 478; Conservative 0; Mismatches 410; Indels 18; Gaps 3;

QY 863 TTCCATGCTGTAGATGATGCTGAAATAGACGCGCTTCCATTTTGATGCTGGGTTTC 922
Db 816 TTCTCTGTTGCTGATGATGATGACATGAATGGCAGCGATTGCACCTGCAAGGTTTC 875
QY 923 CTCAGATGATTCACATGTTCAATACGGAATTCCTGATTCCTCAAGCTGCCCAACAGGAGC 982
Db 876 GATCTATGATGACAAATTTAATGTAAACAATCTCTGCAACAGCTAAAGAGCTTACAAGCTA 935
QY 983 TCGCAGCACAGCTAGACGAGCGAAAGCGCTGGAGATGACAGTCTGCTGCGAGCGCTGG 1042
Db 936 TGGAGGCTCAGCTGA---CTGCGATGTCAGATCACTAGTGTGGTGGGATGGCAGCTCC 992
QY 1043 CAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAAGCTGGGCAACAAC----- 1093
Db 993 CAGCGGAAATACAGCAANTCAAAGATGCTCTTGGCAAGCTTTGAACAACCATCAGCAG 1052
QY 1094 AGGCATACTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGGAGTTC 1153
Db 1053 ATGGTTTGGCTACAGCTATGGGCAAGTGGCTTTTTCAGCTGCCAAGGTTGGAGGAGCT 1112
QY 1154 CTCCGCTGCGACAGAGTTCTATAGGCTCATCTGTAAACAGCTTTTACAAGACCTCAAAAT 1213
Db 1113 CGCAGGACAGCTGGCACTGTCCAGATGAATGTAAACAGCTTTACAAGACAGGTTTTT 1172
QY 1214 CTACAGGTTCT-----GATTATAAACAACAGATATACAGAGTTTATGATGTTACAAAT 1267
Db 1173 CTTGCACTTCTCCAGCTCTTATGACGAGCACTTTCCGATGATATTCTGCTTACAAA 1232
QY 1268 CCATCAATGATGCTTATGGTAGGCGACGAAATGATGCCACTCGTGTATGTATAACAATG 1327
Db 1233 CACTGAATCTTTATATTCCGAAAGCAGAGCGGGTGCAGTCAGCTATTAGTCAAACTG 1292
QY 1328 TAAGTACCCCGCTCTCACAGATCCGTTCTCTAGAGCAGCAAGCTCGAGGACCAAG 1387

Search completed: March 23, 2004, 19:25:15
Job time : 646.252 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:19 ; Search time 4855.74 Seconds

(without alignments)
11389.568 Million cell updates/sec

Title: US-10-608-559-3

Perfect score: 1852

Sequence: 1 atgagtgcgcagataagct.....gatcttcaataataaaga 1852

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estm.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	52.8	2.9	843	28	BZ643413	BZ643413 OGROW57TM
c 2	48.4	2.6	1191	14	CD500670	CD500670 GDA46-D01
c 3	48	2.6	872	14	CK097864	CK097864 UB64CFC02
4	46.6	2.5	359	9	AL840303	AL840303 AL840303

c 5	46.2	2.5	598	14	CD661508
c 6	45.6	2.5	1201	13	BX381961
c 7	45.4	2.5	528	12	BJ340195
c 8	45.4	2.5	724	12	BJ344100
c 9	45.2	2.4	336	14	CB098017
c 10	45.2	2.4	817	14	CD656959
c 11	45.2	2.4	856	13	BUI75987
c 12	45.2	2.4	890	13	BQ424404
c 13	44.8	2.4	484	9	AA314486
c 14	44.8	2.4	885	13	BX425603
c 15	44.4	2.4	856	28	BH132709
c 16	44	2.4	1869	29	AY406745
c 17	43.8	2.4	745	12	BJ430078
c 18	43.6	2.4	701	12	BJ372526
c 19	43.6	2.4	735	12	BJ398404
c 20	43.4	2.3	443	29	FR0008252
c 21	43.4	2.3	801	28	BH314010
c 22	43.4	2.3	850	9	AV406016
c 23	43.4	2.3	861	28	AZ682538
c 24	43.4	2.3	896	12	BP126288
c 25	43.4	2.3	906	28	AZ550193
c 26	43.4	2.3	963	12	BP126164
c 27	43.2	2.3	404	14	CF587772
c 28	43	2.3	420	14	CB798245
c 29	43	2.3	578	13	BU709558
c 30	43	2.3	711	14	CB055699
c 31	43	2.3	756	14	CA326603
c 32	43	2.3	825	14	CF748551
c 33	43	2.3	931	13	BQ900093
c 34	43	2.3	938	13	BU559883
c 35	42.8	2.3	606	12	BJ314050
c 36	42.8	2.3	616	28	AZ525704
c 37	42.8	2.3	764	28	CC017401
c 38	42.8	2.3	1124	13	BX436282
c 39	42.6	2.3	473	12	BI896076
c 40	42.6	2.3	536	14	CD569533
c 41	42.6	2.3	635	13	BW056885
c 42	42.4	2.3	656	12	BQ10829
c 43	42.4	2.3	701	12	BJ490326
c 44	42.4	2.3	980	29	CNS01PHD
c 45	42.2	2.3	843	29	CC569909

ALIGNMENTS

RESULT 1
BZ643413/c
LOCUS BZ643413 843 bp DNA linear GSS 29-JAN-2003
DEFINITION OGROW57TM_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBMA011J17,
Genomic survey sequence.
ACCESSION BZ643413
VERSION BZ643413.1 GI:28104915
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 843)
AUTHORS Whitelaw C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
TITLE Unpublished (2002)
JOURNAL Other GSSs: OGAW57C
COMMENT Contact: Cathy Whitelaw
TIGR Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR

```

Class: sheared ends.
FEATURES             Location/Qualifiers
     source            1..843
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /strain="B73"
                        /db_xref="taxon:4577"
                        /clone="ZMBM0111J17"
                        /clone_lib="ZM 0.7 1.5 kb"
                        /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match          2.9%; Score 52.8; DB 28; Length 843;
Best Local Similarity 49.3%; Pred. No. 0.0048;
Matches 138; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 22 GGTATTGCTTCAGTAACAGCTCGTCTTCTTCTAGCAGATCTGCACAGCTGGACTCAACG 81
Db 702 GCTACTACTACTGCTACTACTGCTACTACTACTGCTACTGCTGCTACTACTACT 643

QY 82 ACAGCGACCGCAGCTAGCGCTCTCCACCCAGCTTTGATGATTATTAAGACTCAAGCGCA 141
Db 642 ACTGCTACTGCTACTGCTGCTACTGCTGCTACTGCTGCTACTGCTACTGCTACT 583

QY 142 ACAGCTTAAGACTACTCTTTACTCTCAACATCACTAGCTGACATACAGAGCTGCTTTGGTG 201
Db 582 ACTGCTGTTGCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTACTACT 523

QY 202 AGCTCCAGATGCTGCTCACTAATTAAGGATACAGCGGCTACTGATAGGAAACGCA 261
Db 522 GGTGCTACTGCTGCTACTGCTACTACTGCTGCTACTGCTGCTGCTGCTGCTGCT 463

QY 262 ATCGCTGGCGAGGCGGAACTAAGAAATGCGGATGCGAGTTA 301
Db 462 ACTGCTACTGCTACTGCTACTACTACTACTGCTGCTGCTGCTA 423

RESULT 2
CD500670/c          1191 bp  mRNA  linear  EST 12-JUN-2003
LOCUS               CDA46-D01.yld-s SHGC-CDA Gasterosteus aculeatus cDNA clone
DEFINITION          CDA46-D01 3', mRNA sequence.
ACCESSION           CD500670
VERSION             CD500670.1 GI:31427701
KEYWORDS            EST.
SOURCE              Gasterosteus aculeatus (three spined stickleback)
ORGANISM            Gasterosteus aculeatus
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                    Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
                    Gasterosteidae; Gasterosteus.
REFERENCE           1 (bases 1 to 1191)
AUTHORS            Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickinson, M.,
                    Schmutz, J., and Myers, R.M.
Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
Contact: Kingsley, DM
HHMI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cmgm.stanford.edu
Plate: 46
High quality sequence start: 12
High quality sequence stop: 866.

FEATURES             Location/Qualifiers
     source            1..1191
                        /organism="Gasterosteus aculeatus"
                        /mol_type="mRNA"
                        /strain="Salinas river, CA"
                        /db_xref="taxon:69293"

```

```

/clone="CDA46-D01"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
/clone_lib="SHGC-CDA"
/note="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI
(5' adaptor); Site 2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dt sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of pBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual pBK-CMV phagemid clones for
EST sequencing."

ORIGIN
Query Match          2.6%; Score 48.4; DB 14; Length 1191;
Best Local Similarity 48.5%; Pred. No. 0.1;
Matches 133; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 371 GPAAACTGACTTCCTTCGACCTTTACAGGCTGCTCTTCCAAATCTGTAGCAACAATA 430
Db 858 GTGATATTCTGACATTAGCCGATAAAGTCTGCAAGACAGAGACTGAAGCAGTGCAG 799

QY 431 ACAAGCAGCTGAGCTTCTTAAGAGATGCAAGTACCCAGTAGTCCAGGGAACGC 490
Db 798 CCACAGATATCAAGCTTTTAAGACGAGATGGAGAGAGATGAAGAGGACGATATTC 739

QY 491 CTGCAATTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAGATAGAGAAG 550
Db 738 CCACATTTGTACATTTCTACAGATCAAAACCATCTGTATAAGTGTGTAAGAGAAAG 679

QY 551 ATGGAATCGGATTAGGANGCATATTTTGAGGACAGACCGCTAGTGAGCTGTAGAAA 610
Db 678 AGGAAAAAGAGATGGGGGGAAGAGAGAGAGAAATCAATAGAGAAGAAAGTA 619

QY 611 ATGCTAAATCTTAACAGTATAAGCAACATAGA 644
Db 618 TTGTAAGAGTGATACGCTGCTGCTACATAAAGA 585

RESULT 3
CD507864            872 bp  mRNA  linear  EST 01-DEC-2003
LOCUS               UB64CPC02.3PR Populus active cambium cDNA library Populus tremula
DEFINITION          cDNA clone UB64CPC02 3', mRNA sequence.
ACCESSION           CK097864
VERSION             CK097864.1 GI:38582189
KEYWORDS            EST.
SOURCE              Populus tremula
ORGANISM            Populus tremula
                    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                    rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE           1 (bases 1 to 872)
AUTHORS            Sterky, F., Bhalarao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,
                    Brunner, A.M., Campa, L., Jonsson-Lindvall, J., Tandré, K.,
                    Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalarao, R.P.,
                    Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J., and Jansson, S.
                    A Populus EST resource for functional genomics
Unpublished (2003)
Other ESTs: UB64CPC02
Contact: Bo Segerman
Umea Plant Science Center, Department of Plant Physiology
Umea University
901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676

```

FEATURES
 source
 Email: bo.segerman@plantphys.umu.se.
 Location/Qualifiers
 1..872
 /organism="Populus tremula"
 /mol_type="mRNA"
 /db_xref="taxon:113636"
 /clone_ref="UB64CPC02"
 /tissue_type="Active cambium"
 /clone_lib="Populus active cambium cDNA library"

ORIGIN

Query Match 2.6%; Score 48; DB 14; Length 872;
 Best Local Similarity 50.4%; Pred. No. 0.11; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
 QY 957 TGATTTCAAGCTGCCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGCAGCAAGCGCTGG 1016
 Db 513 TGTCTGCTGCTGCTGCTCAATTGTTGGAAGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGC 572
 QY 1017 AGATGACAGTCTGCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAACGGCTTAGG 1076
 Db 573 TGCTGCAATTGAGGTAGCTGCTGTGGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGCTGC 632
 QY 1077 TAAAGCTGGGCAACAAAGCGCATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGT 1136
 Db 633 TGTGGCTGCTGCTGTTGAAGCTGTGAGAGCTGCTGCTGCTGCTGCTGCTGTTGAAGC 692
 QY 1137 TGTGAGCGCAGGAGTCTCTCCCGCTGCGCAGCAAGTTCTATAGGGTCATCTGTA 1188
 Db 693 TGTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGTA 744

RESULT 4

AL840303
 LOCUS
 DEFINITION AL840303 F000C Takifugu rubripes cDNA clone F000C10aC11, mRNA
 sequence.
 ACCESSION AL840303
 VERSION AL840303.1 GI:21892241
 KEYWORDS EST.
 SOURCE Takifugu rubripes (Fugu rubripes)
 ORGANISM Takifugu rubripes

REFERENCE

AUTHORS Clark, M.S.
 TITLE Takifugu rubripes ESTs
 JOURNAL Unpublished (2002)
 COMMENT Contact: Clark MS
 MRC Human Genome Mapping Project Resource Centre
 Hinxton, Cambridge, CB10 1SB, UK
 Email: biohelp@hgmrc.mrc.ac.uk
 Vector: pME18S-FL3
 V_type: phagemid
 PRIMER: ME-735FW
 Library created by Koichi Kawakami, Masahide Sasaki, Yutaka Suzuki,
 Sumio Sugano
 The Institute of Medical Science, The University of Tokyo,
 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
 and
 Kiyoshi Kikuchi, Shugo Watabe
 Laboratory of Aquatic Molecular Biology and Biotechnology, Graduate
 School of Agricultural and Life Sciences, The University of Tokyo,
 Bunkyo-ku, Tokyo 108-8639, Japan
 Library sequenced by Melody S. Clark and Amanda Thompson MRC Human
 Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10
 1SB, UK.

FEATURES

source
 Location/Qualifiers
 1..359
 /organism="Takifugu rubripes"
 /mol_type="mRNA"

/db_xref="taxon:31033"
 /clone="F000C10aC11"
 /sex="female"
 /tissue_type="ovary"
 /dev_stage="adult"
 /clone_lib="F000C"
 /note="Vector: pME18S-FL3"

ORIGIN

Query Match 2.5%; Score 46.6; DB 9; Length 359;
 Best Local Similarity 49.8%; Pred. No. 0.16;
 Matches 118; Conservative 0; Mismatches 119; Indels 0; Gaps 0;
 QY 970 GCCCAACAGAGCTCCAGCACAAAGCTAGAGCAGCAAGCGCTGAGATGACAGTGTCT 1029
 Db 47 GCGAAGGAGCAGAGAGCTGCTGAAGCTGCGAGGAGCAGAGCTGCTGAAGCTGTGAAGCT 106
 QY 1030 GCTGCAGCGCTGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGTTAAAGCTGGCAA 1089
 Db 107 GCAGGAGCAGAGCTGCAGGAGCAGAGCTGCAGGAGCAGAGCTGCTGAAGCTGCAGGA 166
 QY 1090 CAACAGGGCATATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGAGCGCAGGA 1149
 Db 167 GCAGAAAGCTGCAGAGCTGCAGGAGCAGAGCTGCTGAAGCTGCTGCAGCAGAGCTGCA 226
 QY 1150 GTTCTCTCCCGCTGCAGCAAGTTCTATAGGTCATCTGTRAAACAGCTTTTCAAGACC 1206
 Db 227 GGAGCAGAGCTGCTGAAGTGCCGGAGCAGAGCTGCTGAAGTCTTGAAGGGTCC 283

RESULT 5

CD661508/c
 LOCUS
 DEFINITION CD661508 598 bp mRNA linear EST 19-JUN-2003
 cDNA 5' similar to TR:Q9UPM8 Q9UPM8 EPSILON-ADAPTIN.1; mRNA
 sequence.
 ACCESSION CD661508
 VERSION CD661508.1 GI:31997248
 KEYWORDS EST.
 SOURCE Eimeria tenella
 ORGANISM Eimeria tenella
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
 Eimeria.

REFERENCE

AUTHORS 1 (bases 1 to 598)
 Liberators, P., Diaz, C., Tang, K., Marra, M., Hillier, L., Kucaba, T.,
 Martin, J., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Florence, N., Shin, I., Jackson, Y.,
 Cardenas, M., McCann, R., Waterston, R., Wilson, R. and Sibley, D.
 WashU-Merck Eimeria tenella project
 TITLE Unpublished (1999)
 JOURNAL Contact: David Sibley, Ph.D.
 COMMENT WashU-Merck Eimeria tenella project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Contact David Sibley (toxoeat@borcim.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco.

FEATURES

source
 Location/Qualifiers
 1..598
 /organism="Eimeria tenella"
 /mol_type="mRNA"
 /db_xref="taxon:5802"
 /dev_stage="LS18"
 /lab_host="E.coli DH10B (GeneHog, Invitrogen, Inc)"
 /clone_lib="Eimeria tenella S5-2 Excised cDNA"
 /note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Sporozoites were obtained from in vitro sporulated and excysted oocysts of E. tenella grown in chickens. cDNA"

was synthesized from poly mRNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA and products were size-selected on sephacryl S500. The cDNA were ligated to EcoRI/XhoI prepared lambda ZapII (Stratagene). The primary library was mass excised using Exasist helper phage (Stratagene). The phagemids were precipitated with PEG 8000, extracted with phenol/chloroform and electroporated into DH10B cells. The library may contain a small percentage of host or bacterial contaminants."

ORIGIN

Query Match 2.5%; Score 46.2; DB 14; Length 598;
Best Local Similarity 49.0%; Pred. No. 0.28;
Matches 123; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 916 GGGTTTCGTGATGATTCACATGTTCAATAGCGAAATCTGATTTCTCAAGTCGCCAA 975
DB 343 GGGTTTAGGGTTATGATGAGATAGGCGCGTGGCTAAACCCCTAGGCTTCTAAA 284

QY 976 CAGGAGCTCGCAGCACAACTAGAGCAGCGAAGCGCTGGAGATGACAGTGTGTGCA 1035
DB 283 CCCAGGTTTCGTTGCTGCTGCTCCAGAGGCTTCAGCCGACGACGCCGCT 224

QY 1036 GCGCTGGCAGATGCTCAGAAAGTTTAGAAGCGGCTCTAGGTAAGCTGGGCAACACAG 1095
DB 223 GCAGCGGCTGCTGTTCCGGCAGTTGAAGATGCTGTTCTATTGCTGCTGTGCTGTC 164

QY 1096 GGCATACCTCAATGCTTAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1155
DB 163 CTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104

QY 1156 CCGCTGCGAGC 1166
DB 103 GCTGCTGCGGC 93

RESULT 6

EX381961
LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSQD1072VF05 3-PRIME, mRNA sequence.

ACCESSION EX381961
VERSION EX381961.1 GI:30453007
SOURCE EST.
ORGANISM Homo sapiens (human)

REFERENCE 1. (bases 1 to 1201)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSQD1072CC03NP1.

FEATURES

Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSQD1072VF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo(dn) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 2.5%; Score 45.6; DB 13; Length 1201;
Best Local Similarity 4.3%; Pred. No. 0.62;
Matches 30; Conservative 243; Mismatches 430; Indels 0; Gaps 0;

QY 740 TGGTAATACAGCGCTGAGAAAGATCTTAAATAATCAACCTGCAGATGGTTCTGATGTC 799
DB 391 KKKKACANNNNNKCKMMNNNNKKKKKCMNNKKNNKKNNKKKKKKNNNNNNNA 450

QY 800 CAAATCCAGGAACACAGTTCGAGGCTCCAAGCAACAGGAAGTAGTATTGCTAGTATTC 859
DB 451 NMMNNKKNN 510

QY 860 GTGTTTCCATGCTGTTAGATGATGCTGAAATGAGACCGCTTCATTTTGTGTTGCTGGT 919
DB 511 CKMMNNKKMMNN 570

QY 920 TTCGTGATGATTCACATGTTCAATACGGAATTCCTGATTTCTCAAGCTGCCAACAGG 979
DB 571 NNN 630

QY 980 AGCTCGCAGCAACAGCTAGAGCAGCGAAGCGCTGGAGATGACAGTGTGCTGCGAGCG 1039
DB 631 KNN 690

QY 1040 TGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAGCTGGGCAACACAGGCA 1099
DB 691 NNN 750

QY 1100 TACTCAATGCTTAGGACAGATCGCTTCTGCTGCTGTTGTGAGCGCAGAGTTCCTCCG 1159
DB 751 NNN 810

QY 1160 CTGACGCAAGTTCATATAGGCTCATCTGTAAGACAGCTTTTACAGACCTCAAAATCTACAG 1219
DB 811 BKGMVCMCKMMNNKKCMCKMMCKMMCKMMCKMMCKMMCKMMCKMMCKMMCKMMCKMM 870

QY 1220 GTTCTGATTATAAACAAGATATGAGAGGTTATGATGCTTACAAATCCATCAATGATG 1279
DB 871 KDNMMNN 930

QY 1280 CCTATGGTAGGCGACGAAATGATGCGATGCTGATGATGATAAACAATGTAAGTACCCCG 1339
DB 931 KKKMMNN 990

QY 1340 CTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGAGCTCGAGACAGAAAAAAGATC 1399
DB 991 MCMCKKKBKMKVCKCKKCMCKMMCKMMCKMMCKMMCKMMCKMMCKMMCKMMCKMMCK 1050

QY 1400 AAGCCCTCGCTAGGCTGATTTCTGCGCAATAGCAGAACTCTTGG 1442
DB 1051 KKKKCKKKBKMMCKMMNN 1093

RESULT 7

EX340195/c
LOCUS BJ340195 Dictyostelium discoideum cDNA library, AF Dictyostelium
DEFINITION dictyostelium cDNA clone dda13f05 3', mRNA sequence.

ACCESSION BJ340195
VERSION BJ340195.1 GI:19248557
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

REFERENCE 1. (bases 1 to 528)
Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.
Full length cDNA of Dictyostelium discoideum at the aggregation stage
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information

National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES

source
1. .528
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddal2f05"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_lib="Dictyostelium discoideum cDNA library, AF"

ORIGIN

Query Match 2.5%; Score 45.4; DB 12; Length 528;
Best Local Similarity 44.8%; Pred. No. 0.43; 216; Indels 0; Gaps 0;
Matches 175; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 514 GATCAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGGATTAGGGATGCA 573
DB 480 GATGCCAGAAAGCTGCTGATGCAGAGCAAGAAAGCTGCAGAGCAAGAAAGCTGCT 421

QY 574 TATTTTCAGGACAGAAAGCTAGTGGAGCTGTAGAAAATGTAATCTTAATACAGTATA 633
DB 420 GATGCAAGCAAGAAAGCTGCTGATGCCAAGAAAGCTGCTGATGAAGAAAGCA 361

QY 634 AGCAACATAGATTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 693
DB 360 AAGAAAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 301

QY 694 CAGAAAAAGTTCCCGAGCTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAATACAGCT 753
DB 300 GCAAGAAAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 241

QY 754 GAGAAAGATCTTAAATAATCAACCTGCAGATGTTCTGATGTTCCAAATCCGGAAGCT 813
DB 240 GCTAAGAAAGCTGCTGATGAAGAAAGCAAGAAAGCTGCTGATGCTGATGCTGATGCT 181

QY 814 ACAGTTGAGGCTCCAGCAACAGGAAGTAGTATTGGTAGTATTGGTATTCGTTCCATGCTG 873
DB 180 GAAGAAAGAAAGCAAGAAAGCTGCTGATATCAAGAAAGCTGCAGAAAGATGCCAAAAA 121

QY 874 TTAGATGATGCTGAAAAATGAGACCGCTTCCA 904
DB 120 GCAGAGATGCCAAAAAGCAGAGATGCCA 90

RESULT 8
BU344100/c 724 bp mRNA linear EST 06-MAR-2002
LOCUS BU344100 Dictyostelium discoideum cDNA library, AF Dictyostelium
DEFINITION discoideum cDNA clone ddal2f03, mRNA sequence.

ACCESSION BU344100
VERSION BU344100.1 GI:19214607
KEYWORDS EST.

SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Metazoa; Dictyosteliida; Dictyostelium.

AUTHORS 1 (bases 1 to 724)

TITLE Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

Full length cDNA of Dictyostelium discoideum at the aggregation stage

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

FEATURES

source
1. .724
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddal2f05"
/sex="mat A"
/dev_stage="Aggregation stage"
/clone_lib="Dictyostelium discoideum cDNA library, AF"

ORIGIN

Query Match 2.5%; Score 45.4; DB 12; Length 724;
Best Local Similarity 44.8%; Pred. No. 0.52; 216; Indels 0; Gaps 0;
Matches 175; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 514 GATCAGACAGATGCTACAGCGACACAGATAGAGAAAGATGGAATGGATTAGGGATGCA 573
DB 679 GATGCCAGAAAGCTGCTGATGCAGAGCAAGAAAGCTGCAGAGCAAGAAAGCTGCT 620

QY 574 TATTTTCAGGACAGAAAGCTAGTGGAGCTGTAGAAAATGTAATCTTAATACAGTATA 633
DB 619 GATGCAAGCAAGAAAGCTGCTGATGCCAAGAAAGCTGCTGATGAAGAAAGCA 560

QY 634 AGCAACATAGATTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 693
DB 559 AAGAAAGCTGCTGATGCCAAAAAGCTGCCGATGCCAAGAAAGCTGCTGATGAAGAA 500

QY 694 CAGAAAAAGTTCCCGAGCTCTCCAAATCTTCAAGAAGCGGAACAAATGGTAATACAGCT 753
DB 499 GCAAGAAAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 440

QY 754 GAGAAAGATCTTAAATAATCAACCTGCAGATGTTCTGATGTTCCAAATCCGGAAGCT 813
DB 439 GCTAAGAAAGCTGCTGATGAAGAAAGCAAGAAAGCTGCTGATGCTGATGCTGATGCT 380

QY 814 ACAGTTGAGGCTCCAGCAACAGGAAGTAGTATTGGTAGTATTGGTATTCGTTTCCATGCTG 873
DB 379 GAAGAAAGAAAGCAAGAAAGCTGCTGATATCAAGAAAGCTGCAGAAAGATGCCAAAA 320

QY 874 TTAGATGATGCTGAAAAATGAGACCGCTTCCA 904
DB 319 GCAGAGATGCCAAAAAGCAGAGATGCCA 289

RESULT 9

CB098017

LOCUS CB098017

DEFINITION Ku48b04.y1 Strongyloides ratti PA female naive pAMP1 v1

Strongyloides ratti cDNA 5' similar to contains element XTR

repetitive element ;, mRNA sequence.

ACCESSION CB098017

VERSION CB098017.1 GI:27922200

KEYWORDS EST.

SOURCE Strongyloides ratti

ORGANISM Strongyloides ratti

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

AUTHORS Panagrolaimoidae; Strongyloidea; Strongyloides.

1 (bases 1 to 336)

McCart, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,

Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,

Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,

Tsagaris, V., R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,

Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,

Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

McCart, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

CONTACT: McCarter, J.P.

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy, Brandi Chiapelli and Dr. James McCarter at Washington University, St. Louis. The cDNA library was made by using Dynabead oligo-dT priming (DynaL). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis kit from Clontech. Directionally cloned into the UPG sites of pAMP1. Parasitic adult females were collected from immunologically naive animals and provided by Dr. Mark Viney of Bristol, UK.
Seq primer: -40RP from Gibco
High quality sequence stop: 298.

FEATURES

source

```
1. 336
/organism="Strongyloides ratti"
/mol_type="mRNA"
/db_xref="taxon:34506"
/dev_stage="parasitic adult females"
/lab_host="DH10B"
/clone_lib="Strongyloides ratti PA female naive pAMP1 v1"
/notes="Vector: pAMP1 (Gibco); Site1: NotI; Site2: SalI;
The library was constructed by Claire Murphy, Brandi
Chiapelli and Dr. James McCarter at Washington University,
St. Louis. The cDNA was made by using Dynabead oligo-dT
priming (DynaL). PCR based library using a modified
protocol from the SMART PCR cDNA Synthesis Kit from
Clontech. Directionally cloned into the UDG sites of
pAMP1. Parasitic adult females were collected from naive
animals and provided by Dr. Mark Viney of Bristol, UK."
```

ORIGIN

```
Query Match 2.4%; Score 45.2; DB 14; Length 336;
Best Local Similarity 47.8%; Pred. No. 0.38;
Matches 131; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Qy 544 GAGAAAGATGGAATGCGATTAGGATGATATTTTCAGACAGACAGCGTAGTGAGCT 603
Db 31 GAAACAGAGATATCTGCTTCTTAAGAAACTGAAGTTACAGAAACTTCTGGTACAGAGAA 90

Qy 604 GTAGAAATGTAATCTAATTAACAGTATAGCAACATAGATCAGTAAAGCAGCAATC 663
Db 91 ACAGAAATGTCGAATCTAAGAAACTGAATTCCTGAATCTTCTGGTACAGAAAGC 150

Qy 664 GCTACTCTAGACACAAATAGCTGAAGCTCAGAAAAAGTCCCGGACTCTCCAAATCTTT 723
Db 151 GAATTCATATCTGACAAACAGAAACTCTGTTACATATCCAGATATTCAGGTAGC 210

Qy 724 CAAGAAGCGGAACAAATGGTAAATACAGCTGAGAAAGATCTTAAANAATCAACACCTGCA 783
Db 211 ACTGAATCTAGAGAAACAGAAAGTATCTGCTTCTTAAGAAACTGAAGTCACTGAAACTTCT 270

Qy 784 GATGGTTCTGATGTTCCAAATCCAGGAAGTACAG 817
Db 271 GCCACAGAGAAACTGAATCTGTAATCTTAAG 304
```

RESULT 10

```
CD656959/c
LOCUS CD656959 817 bp mRNA linear EST 18-JUN-2003
DEFINITION AGENCOURT_14540498 NIA Human H1 Embryonic Stem Cell cDNA Library
(Long) Homo sapiens cDNA clone IMAGE:30421464 5', mRNA sequence.
ACCESSION CD656959
VERSION CD656959.1 GI:31898704
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 817)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
```

National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbbs-remail.nih.gov
Tissue Procurement: Irene Genis and Mahendra Rao, NIA
cDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c lone distribution information
can be found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: NDAM499 row: f column: 01
High quality sequence stop: 760.
Location/Qualifiers

FEATURES

source

```
1. 817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30421464"
/tissue_type="Embryonic stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library
(Long)"
/notes="Vector: pCMV-Sport6; Site1: NotI; Site2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from WiCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GAP, and EOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/Gibco) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with Trizol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:11544199] Double-stranded cDNAs were
synthesized with an oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker LL-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-SPORT6 plasmid vector. The average insert size is
about 3.6kb."
```

ORIGIN

```
Query Match 2.4%; Score 45.2; DB 14; Length 817;
Best Local Similarity 45.3%; Pred. No. 0.64;
Matches 164; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

Qy 954 TCTGATCTCAAGCTGCCCAACAGAGCTCGCAGACACAGCTAGACGACGAAACCGC 1013
Db 571 TCTGGAATCTCCACTGCTGGATCATGTCCAAAAAGCTGCTGCTGAGCGATGTGGT 512

Qy 1014 TGGAGATGACAGTGTGCTGCAGCGTGGCAGATGTCAGAAAGCTTTAGAACGGCTCT 1073
Db 511 GGCTGCTGTTGCTGCTGCTGCTCGGGGGCTGCTGAGTGGACCTGATGCTCAGGGCC 452

Qy 1074 AGGTAAGACTGGGCACACACAGGGGCATCTCAATGCTTTTAGGACAGATCGCTTCTGTCG 1133
Db 451 TGTGACCTCTGACTCAAGTGTGGCTGTCTGACTTTTGATCTGCTGCTGCTGCTGTCG 392

Qy 1134 TGTGTGAGCGCAGAGGAGTCTCCCGCTGCAGCAAGTCTCTATAGGTCATCTGTAAACA 1193
```

```

Db      391  TGGTGCATCGCTGGAGCTGCTGCTGAATCTGCTGTTGCTGCTGCTGGAGCAGG 332
QY      1194  GCCTTACAGACCTCAAAATCTACAGCTTCTGATATATAAAACACAGATATCAGCAGGTTA 1253
Db      331  CTGTGCTGCTGCTGGAGCTGGCAAGCTGCTGCTGCTGGAGCTGGTCTGCTGCAGCTGA 272
QY      1254  TGAATGCTTACAATCCATCAATGATGCTATGGTGGGACAGAAATGATGCGACTGCTGA 1313
Db      271  TGAACCTGCTGCTGCAGGGCATGCTGCTGCTGAAAGGCTGGAGCTGCTGCTGAGGGCGA 212
QY      1314  TG 1315
Db      211  TG 210

RESULT 11
LOCUS   BUI75987              856 bp      mRNA      linear      EST 04-SEP-2002
DEFINITION   AGENCOURT_7782344 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6137044
5', mRNA sequence.
BUI75987
VERSION      BUI75987.1 GI:22689971
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 856)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            plate: LLAM13450 row: d column: 05
            High quality sequence stop: 602.
FEATURES             Location/Qualifiers
     source           1..856
     organism="Homo sapiens"
     mol_type="mRNA"
     db_xref="taxon:9606"
     clone="IMAGE:6137044"
     tissue_type="retinoblastoma"
     lab_host="DH10B (phage-resistant)"
     clone_lib="NIH_MGC_67"
     note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
     Average insert size 1.75 kb. Library constructed by Life
     Technologies."

ORIGIN
Query Match      2.4%; Score 45.2; DB 13; Length 856;
Best Local Similarity 45.3%; Pred. No. 0.66;
Matches 164; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY      954  TCTGATTCATGCTGCCACAGAGCTCGACACAGCTAGACGACGAAAGCGC 1013
Db      597  TCTGGAATCTCACTGCTGGATCATGTCCAAAAGCTGATGCTGCTGAGGCGATGTTGT 538
QY      1014  TGGAGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAGCGGCTCT 1073
Db      537  GCGTGTCTGCTGCTGCTGCTGCGGGGCTCTGAGGTGGAACTGATGCTGCGAGCGCC 478
QY      1074  AGGTAAGCTGGGCAACAACAGGGCATACTCAATGCTTTAGGACAGATGCTTCTGCTGC 1433
Db      477  TGTGACGTCTGACTCAAGTGTGGCGCTGCTGACTTTGATCTGCTGCTGCTGCTGC 418

```

```

QY      1134  TGTGTGAGCGGAGGAGTTCTCCCTGCTGAGCAAGCTTCTATAGGTCATCTCTAAAACA 1193
Db      417  TGTGTGATCGCTGGAGCTGCTGCTGCTGAATCTGCTGTTGCTGCTGCTGGAGCAGG 358
QY      1194  GCCTTACAGACCTCAAAATCTACAGCTTCTGATATATAAAACACAGATATCAGCAGTTA 1253
Db      357  CTGTGCTGCTGCTGGAGCTGGGCGAGCTGCTGCTGCTGGAGCTGGTCTGCTGCAGCTGA 298
QY      1254  TGAATGCTTACAATCCATCAATGATGCTATGGTGGGACAGAAATGATGCGACTGCTGA 1313
Db      297  TGAACCTGCTGCTGCAGGGCATGCTGCTGCTGAAAGGCTGGAGCTGCTGCTGAGGGCGA 238
QY      1314  TG 1315
Db      237  TG 236

RESULT 12
LOCUS   BQ424404/c              890 bp      mRNA      linear      EST 23-MAY-2002
DEFINITION   AGENCOURT_7907155 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154729
5', mRNA sequence.
BQ424404
VERSION      BQ424404.1 GI:21119719
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 890)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            plate: LLAM13496 row: e column: 02
            High quality sequence stop: 673.
FEATURES             Location/Qualifiers
     source           1..890
     organism="Homo sapiens"
     mol_type="mRNA"
     db_xref="taxon:9606"
     clone="IMAGE:6154729"
     tissue_type="retinoblastoma"
     lab_host="DH10B (phage-resistant)"
     clone_lib="NIH_MGC_67"
     note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
     Average insert size 1.75 kb. Library constructed by Life
     Technologies."

ORIGIN
Query Match      2.4%; Score 45.2; DB 13; Length 890;
Best Local Similarity 45.3%; Pred. No. 0.67;
Matches 164; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY      954  TCTGATTCATGCTGCCACAGAGCTCGACACAGCTAGACGACGAAAGCGC 1013
Db      691  TCTGGAATCTCCACTGCTGGATCATGTCCAAAAGCTGATGCTGCTGAGGCGATGTTGT 622
QY      1014  TGGAGATGACAGTGTGCTGCGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAGCGGCTCT 1073
Db      621  GCGTGTCTTGTGCTGCTGCTGCTGCGGGGCTGCTGAGGTGGAACTGATGCTGCGAGCGCC 562
QY      1074  AGGTAAGCTGGGCAACAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTGC 1133

```

```
/cell_type="KM12C"
/cell_line="KM12C(HCC)-parental human colon
carcinoma/Dukes B2"
/clone_lib="Colon carcinoma (HCC) cell line II"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN
Query Match      2.4%; Score 44.8; DB 9; Length 484;
Best Local Similarity 51.5%; Pred. No. 0.61;
Matches 103; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 967 GCTGCCCAACAGGAGCTGCAGCACAAGCTAGACAGCGAAGCGCTGGAGATGACAGT 1026
Db 379 GCTGCTGAGCTGGCGGAGCTGCTGCTGGAGCTGGTGTGCTGCAGCTGATGGAAC 320

QY 1027 GCTGCTGAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAGCTGGG 1086
Db 319 GCTGCTGAGGCGATGCTGCTGCTGAAAGGCTGGAGCTGCTGAGGCGGATGCACT 260

QY 1087 CAACAACAGGCGATCAATGCTTTAGACAGATCGCTTCTGCTGTTGTGAGCGCA 1146
Db 259 GCTGTGGAGATGCAACTGTTGCTGTGAAACTGTAGCTGTTGCTGTGAAAAGCTATGCG 200

QY 1147 GGAGTTCTCCCGCTGCAGC 1166
Db 199 GATGGCTTGTGCTGCTGC 180

RESULT 14
BX425603 885 bp mRNA linear EST 15-MAY-2003
BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB022ZA07 3-PRIME, mRNA sequence.
ACCESSION BX425603
VERSION BX425603.1 GI:30770486
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 885)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBB022ZA07FP1.
FEATURES
source
1..885
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match      2.4%; Score 44.8; DB 13; Length 885;
Best Local Similarity 4.1%; Pred. No. 0.87;
Matches 25; Conservative 165; Mismatches 413; Indels 0; Gaps 0;

QY 962 CTCAGAGTGCACAGAGAGCTGCGACACACAGTAGACGAGCAAGCTAGAGCGAAGCGCTGGAGATG 1021
```


Search completed: March 24, 2004, 05:56:10
Job time : 4861.4 secs

RESULT 15
 BH132709/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

856 bp DNA linear GSS 07-AUG-2001
 ENTING59FF Entamoeba histolytica Sheared DNA Entamoeba histolytica
 genomic, genomic survey sequence.
 BH132709 GI:15091760
 GSS.
 Entamoeba histolytica
 Entamoeba histolytica
 Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 1 (bases 1 to 856)
 Loftus,B., Wang Z., Van Aken,S. and Fraser,C.
 Determination of clone end sequences from Entamoeba histolytica
 HM1:IMSS sheared DNA library (2001)
 Unpublished (2001)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: biloftus@tigr.org
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
 DNA library
 Seq primer: M13-Forward
 Class: shotgun
 High quality sequence start: 3
 High quality sequence stop: 856.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2004, 13:27:15 ; Search time 7197.51 Seconds
(without alignments)
11152.641 Million cell updates/sec

Title: US-10-608-559-3
Perfect score: 1852
Sequence: 1 atgagttggcagataagct.....gatcttcataataaaaga 1852

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_in.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1852	100.0	1852	6	AX045133	AX045133 Sequence
2	1848.8	99.8	2156	6	AX045131	AX045131 Sequence
3	1848.8	99.8	10421	1	AE001654	AE001654 Chlamydia
4	1848.8	99.8	12548	1	AE002165	AE002165 Chlamydia
5	1848.8	99.8	30066	1	AE017159	AE017159 Chlamydia
6	1848.8	99.8	300550	1	AE002547	AE002547 Chlamydia
7	1836.2	99.1	110000	6	AX318754_08	Continuation (9 of
8	1748.8	94.4	1956	6	AX338293	AX338293 Sequence
9	1748.8	94.4	1956	6	AX349491	AX349491 Sequence
10	1748.8	94.4	1956	6	AX361992	AX361992 Sequence
11	1165	62.9	2148	1	CHT76KDA	L23921 Chlamydia p
12	1161.8	62.7	2238	6	AX045137	AX045137 Sequence
13	1151	62.1	1456	6	AX045135	AX045135 Sequence
14	361.2	19.5	304769	1	AE016994	AE016994 Chlamydia
15	187.2	10.1	1537	6	BD264444	BD264444 Compounds
16	187.2	10.1	1537	6	AR229257	AR229257 Sequence
17	187.2	10.1	1537	6	AR308168	AR308168 Sequence
18	187.2	10.1	1537	6	AR321875	AR321875 Sequence
19	187.2	10.1	1537	6	AX155814	AX155814 Sequence
20	187.2	10.1	1537	6	AX361664	AX361664 Sequence
21	187	10.1	10954	1	AE001333	AE001333 Chlamydia
22	186	10.0	1944	6	AX786284	AX786284 Sequence
23	185.4	10.0	1941	6	AX362019	AX362019 Sequence
24	184	9.9	1171	6	AX338243	AX338243 Sequence
25	184	9.9	1834	6	AX338268	AX338268 Sequence
26	183	9.9	1983	6	AX338258	AX338258 Sequence
27	148.4	8.0	150	6	AR144067	AR144067 Sequence
28	118	6.4	14698	1	AE002357	AE002357 Chlamydia
29	54.4	2.9	7218	6	I66494	I66494 Sequence (4 of
30	52.4	2.8	77835	2	PFMAL13P2_3	Continuation (24 of
31	52.4	2.8	110000	2	PFMAL13_23	Continuation (24 of
32	51.6	2.8	2205	6	AX317736	AX317736 Sequence
33	50.6	2.7	423	1	MC049	Z33043 M. capricolu
34	49.8	2.7	2000	6	AX655393	AX655393 Sequence
35	49.4	2.7	5098	1	MCU51235	U51235 Mycoplasma
36	49	2.6	1990	6	AR268644	AR268644 Sequence
37	49	2.6	2085	6	AR008322	AR008322 Sequence
38	49	2.6	2085	6	AR027306	AR027306 Sequence
39	49	2.6	2085	6	AR035771	AR035771 Sequence
40	49	2.6	2085	6	AR037631	AR037631 Sequence
41	49	2.6	2085	6	AR078816	AR078816 Sequence
42	49	2.6	2085	6	AR079063	AR079063 Sequence
43	49	2.6	2085	6	AR084396	AR084396 Sequence
44	49	2.6	2085	6	I16740	I16740 Sequence 1
45	49	2.6	2085	6	I70472	I70472 Sequence 1

ALIGNMENTS

RESULT 1
AX045133
LOCUS
DEFINITION
Sequence 3 from Patent WO0066739.
ACCESSION
AX045133
VERSION
AX045133.1 GI:11343732
KEYWORDS
Chlamydia pneumoniae
SOURCE
Chlamydia pneumoniae
ORGANISM
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE
1
Murdin, A.D., Oomen, R.P., Wang, J. and Dunn, P.
AUTHORS
i (chlamydia) antigens and corresponding dna fragments and uses
TITLE
Thereof
JOURNAL
Patent: WO 0066739-A 3 09-NOV-2000;

0.996 (99.6%)
PAT 24-NOV-2000

Aventis Pasteur Limited (CA)	
Location/Qualifiers	
1. .1852	
/organism="Chlamydomophila pneumoniae"	
/mol_type="unassigned DNA"	
/db_xref="taxon:83558"	
1. .1752	
/notes="unnamed protein product"	
/codon_start=1	
/transl_table=11	
/protein_id="CAC17378.1"	
/db_xref="GI:111343733"	
/db_xref="REMBL:CAC17378"	
/translation="MSLADKLGIGLASNSSSSSTSRSDVSTTATATPPPTDDYKT	
QAQAYDTITSTSLADIQAALVSLQDAVINIKDITADTDETAIAAEWETKNADAVKY	
GAQITELAKYFISNQAILDLGLKLTSLFDLQAALLOSVAANNKAAELKEMQNDPVVP	
GKTPAIASLVDDTDATQIEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAK	
AAIATATQIAEAKQKPPDSPILQAEQWVIQAEKDKNKIKPADGSPVNPPTVGGSS	
KQGGSSISIRVSMLLDDAENETASIIIMSGFRQIMHMENTENPDSQAQOELAAQARA	
AKAAGDSAAALADAQKALEALGKAGQGGILNALGQASRAVSVAGVPPRAASRI	
GSSVKLYKTSKSTGSDYKTIQISAGYDAYKSINDAYGRANDATRDVINNVSTPALIR	
SVPRATEARPEKTDQALARVIGSNRSRTILGDVYSQVSLQSVNQIITQSNPQANNEI	
RQKLTSVATPKPPQPGYPIVOLSNDSTQKFIKLESIFAEGSRTAAETIKALSPETNSLF	
IQQLVNIIGSLYSYLQ"	
ORIGIN	
Query Match	100.0%; Score 1852; DB 6; Length 1852;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1852; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
DB	1 ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
QY	61 TCTGACAGCTGGGACTCAACGACGCGACCGACCTACGCTCTCTCCACCCAGCTTTGAT 120
DB	61 TCTGACAGCTGGGACTCAACGACGCGACCGACCTACGCTCTCTCCACCCAGCTTTGAT 120
QY	121 GATTATAAGACTCAAGCGCAACAGCTTACGATCTACTTCTTACTCAACATCACTAGCT 180
DB	121 GATTATAAGACTCAAGCGCAACAGCTTACGATCTACTTCTTACTCAACATCACTAGCT 180
QY	181 GACATACAGGCTGCTTGGTGAGCTCCAGGATGCTGTCACATATATAAAGGATACAGCG 240
DB	181 GACATACAGGCTGCTTGGTGAGCTCCAGGATGCTGTCACATATATAAAGGATACAGCG 240
QY	241 GCTACTGATGAGGAAACCGCAATGCTGCGGAGTGGGAAAATAAGATGCGGATGCGAGTT 300
DB	241 GCTACTGATGAGGAAACCGCAATGCTGCGGAGTGGGAAAATAAGATGCGGATGCGAGTT 300
QY	301 AAAAGTTGGCGCGCAATTTACGAAATTTAGCGAAATATGCTTGGATAAACCAAGCGATTCTT 360
DB	301 AAAAGTTGGCGCGCAATTTACGAAATTTAGCGAAATATGCTTGGATAAACCAAGCGATTCTT 360
QY	361 GACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 420
DB	361 GACTCTTTAGGTAAACTGACTCTCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 420
QY	421 GCAACATATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
DB	421 GCAACATATACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
QY	481 GGGAAAACCCCTGCAATCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
DB	481 GGGAAAACCCCTGCAATCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
QY	541 ATAGAGAAAGATGGAAATCGGATAGGGATGATATTTTTCAGGACAGAACCGCTAGTGA 600
DB	541 ATAGAGAAAGATGGAAATCGGATAGGGATGATATTTTTCAGGACAGAACCGCTAGTGA 600
QY	601 GCTGTAGAAATGCTTAAATCTTAATACAGTATACAGCAATAGATTCAGCTAAAGCGCA 660
DB	601 GCTGTAGAAATGCTTAAATCTTAATACAGTATACAGCAATAGATTCAGCTAAAGCGCA 660
QY	661 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAATT 720
DB	661 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAGTTTCCCGACTCTCCAATT 720
QY	721 CTTCAAGAAAGCGGAAACAAATGGTAAATACAGGGCTGAGAAAAGATCTTAAAAATATCAAACT 780
DB	721 CTTCAAGAAAGCGGAAACAAATGGTAAATACAGGGCTGAGAAAAGATCTTAAAAATATCAAACT 780
QY	781 GCAGATGGTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 840
DB	781 GCAGATGGTCTGATGTTCCAAATCCAGGAACTACAGTTGGAGGCTCCAAAGCAACAGGA 840
QY	841 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTCTTATAGATGATGCTCAAAAATAGAGACGCT 900
DB	841 AGTAGTATTGGTAGTATTGCTGTTTCCATGCTCTTATAGATGATGCTCAAAAATAGAGACGCT 900
QY	901 TCCATTTTGTATGCTCGGTTTCTGATGATTCAATGTTTCAATACGGAAAATCCCTGAT 960
DB	901 TCCATTTTGTATGCTCGGTTTCTGATGATTCAATGTTTCAATACGGAAAATCCCTGAT 960
QY	961 TCTCAAGCTGCCCCAACAGGAGCTCGCAGCACAAGCTAGACGCGGAAAGCCGCTGGAGAT 1020
DB	961 TCTCAAGCTGCCCCAACAGGAGCTCGCAGCACAAGCTAGACGCGGAAAGCCGCTGGAGAT 1020
QY	1021 GACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTCTAGGTAAA 1080
DB	1021 GACAGTCTGCTGCGAGCGCTGGCAGATGCTCAGAAAAGCTTTAGAACGCGCTCTAGGTAAA 1080
QY	1081 GCTGGGGCAACAAAGGAGCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
DB	1081 GCTGGGGCAACAAAGGAGCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTTGTG 1140
QY	1141 AGCGCAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAACAGCTTTAC 1200
DB	1141 AGCGCAGGAGTCTCTCCGCTGCGAGCAAGTTCTATAGGTCATCTGTAACAGCTTTAC 1200
QY	1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
DB	1201 AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT 1260
QY	1261 TACAAATCCATCAATGATGCTATGTTAGGCGCAGAAATGATGCGACTCGTGTGATGATA 1320
DB	1261 TACAAATCCATCAATGATGCTATGTTAGGCGCAGAAATGATGCGACTCGTGTGATGATA 1320
QY	1321 AACAACTGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 1380
DB	1321 AACAACTGTAAGTACCCCGCTCTCACACGATCCGTTCTTAGAGCAGCAACAGAAAGCTCGA 1380
QY	1381 GGACCCAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCTTT 1440
DB	1381 GGACCCAGAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGGCAATAGCAGAACTCTTT 1440
QY	1441 GGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTGAATGAGATCAGCTCAGTCCGAAT 1500
DB	1441 GGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTGAATGAGATCAGCTCAGTCCGAAT 1500
QY	1501 CCTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCCCTCCA 1560
DB	1501 CCTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGCAAGTACAAAGCCCTCCA 1560
QY	1561 CAGTTTGGCTATCTTATAGTGCAACTTTTCTTAATGACTCTACACAGAAAGTTCATAGCTAAA 1620
DB	1561 CAGTTTGGCTATCTTATAGTGCAACTTTTCTTAATGACTCTACACAGAAAGTTCATAGCTAAA 1620
QY	1621 TTAGAAAGTTTGTGTTGCTGAGGATCTAGGACAGCAGCTGAAATAAAGACTTTTCTTTT 1680
DB	1621 TTAGAAAGTTTGTGTTGCTGAGGATCTAGGACAGCAGCTGAAATAAAGACTTTTCTTTT 1680
QY	1681 GAAACGAACTCTCTTTTATTTCAGCAGTGTGCTCAATATFCGGCTCTCTATATATTCGGT 1740
DB	1681 GAAACGAACTCTCTTTTATTTCAGCAGTGTGCTCAATATFCGGCTCTCTATATATTCGGT 1740

QY 1741 TATCTCCAAATACACACCTTAAGTGTTCGTTTGGAGAGATTATATATGTGCTTTGGTAGG 1800
DB 1741 TATCTCCAAATACACACCTTAAGTGTTCGTTTGGAGAGATTATATATGTGCTTTGGTAGG 1800
QY 1801 CCTTTGTTGAGCCCTTACCAACACTAGAGATCTTCAATAATAAAGA 1852
DB 1801 CCTTTGTTGAGCCCTTACCAACACTAGAGATCTTCAATAATAAAGA 1852

RESULT 2
AX045131 2156 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 1 from Patent WO0066739.
DEFINITION AX045131
ACCESSION AX045131
VERSION AX045131.1 GI:11343730
KEYWORDS
SOURCE Chlamydomophila pneumoniae
ORGANISM Chlamydomophila pneumoniae
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.

REFERENCE 1
Murdin, A.D., Oomen, R.P., Wang, J. and Dunn, P.
i(chlamydia) antigens and corresponding dna fragments and uses
Thereof
Patent: WO 0066739-A 1 09-NOV-2000;
Aventis Pasteur Limited (CA)
FEATURES
source
1..2156
/organism="Chlamydomophila pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
101..2056
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC17329.1"
/db_xref="GI:11343731"
/db_xref="REMBL:CAC17329"
/translation="MVNPIGPIDETPPADLSAQLEASANKSABEAOITAGAE
AKPKESKTSVERMSILSALNADKIGIASNSSSSSTSRSDVSTTATPT
PPPTDDYKQATYDITFTSLADIQAALVSDQAVTNIKDTAADEETAAEW
KEMADAVKGAQITELAKYADNQAILDLSLGLTSLFDLLOALIQSVANNKNAELL
RTMONPVVPGKTPAIAGSLVDOTDATQIEKGNJRDVAFQGNASGAVENAKSN
NSISNIDAKALATKQIAEAKKFPDPSPILOEAEQMVIOAEKDLNKPAGSDV
PNFTGVGSKQGSIGSIRVMSLLDABNETAGILMSGRFQMLHMENTENPDSQAA
QOELAAQARAAGAAGDSDAALADAKALEAALGAGQGGQILNALGQIAAAVVA
GVPPAAASISGSVQLYKTSKSDYKQISAGYDAYKSINDAYGARNDAITDVI
NVNPTALTRSVPRARTFARGEKTDQALRVISGNSRTLGDVYQVSAALQSVMIQ
SNPOANNEIRKQLTSVTKPPQFGYPPVQLSNDSTQKFIKLESLEFAEGSRTAERIK
ALSFTNSLFIQQLVNLGSLYGLQ"

ORIGIN
Query Match 99.8%; Score 1848.8; DB 6; Length 2156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGTCTGGCAGATAAGTGGGTATGCTTCTAGTAAACAGCTCGCTCTCTACTAGCAGA 60
DB 305 ATGAGTCTGGCAGATAAGTGGGTATGCTTCTAGTAAACAGCTCGCTCTCTACTAGCAGA 364
QY 61 TCTGCAGACGTGGACTCAACGACAGCGACCGACCTACGCTCTCTCCACCCACGTTTGAT 120
DB 365 TCTGCAGACGTGGACTCAACGACAGCGACCGACCTACGCTCTCTCCACCCACGTTTGAT 424
QY 121 GATTATAGACTCAACGCAAAACAGCTTACGATCTATCTTCTCAACATCACTAGCT 180
DB 425 GATTATAGACTCAACGCAAAACAGCTTACGATCTATCTTCTCAACATCACTAGCT 484
QY 181 GACATACAGGCTCTTGTGAGCTCCAGGATGCTGCTCACTAATAAAGATACAGCG 240
DB 485 GACATACAGGCTCTTGTGAGCTCCAGGATGCTGCTCACTAATAAAGATACAGCG 544
QY 241 GCTACTGATGAGAAACCGCAATCGCTGCGGAGTGGAAACTAAGAAATGCCGATCGAGTT 300

```
QY 1381 GGACCAAGAAAAAAGATCAAGCCCTCGTAGGTGATTTCTGCGCAATAGCAGAACTCTT 1440
Db 1685 GGACCAAGAAAAAAGATCAAGCCCTCGTAGGTGATTTCTGCGCAATAGCAGAACTCTT 1744
QY 1441 GGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTGTAATGCGAGATCATCCAGTCGAAT 1500
Db 1745 GGAGATGCTATAGTCAAGTTTCGGCACTCAATCTCTGTAATGCGAGATCATCCAGTCGAAT 1804
QY 1501 CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGCGAGTACAAAGCCCTCCA 1560
Db 1805 CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGCGAGTACAAAGCCCTCCA 1864
QY 1561 CAGTTTGGCTATCCTTATGTCACATTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1620
Db 1865 CAGTTTGGCTATCCTTATGTCACATTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1924
QY 1621 TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAATAAAGACACTTTCCTTT 1680
Db 1925 TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAATAAAGACACTTTCCTTT 1984
QY 1681 GAAACGAACTCCTTGTATTATCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTCGTT 1740
Db 1985 GAAACGAACTCCTTGTATTATCAGCAGGTGCTGCTCAATATCGGCTCTCTATATTCGTT 2044
QY 1741 TATCTCCAAATCAACACTAAGTCTGTTGAGAGATTTATGATGCTTTCGTAAGG 1800
Db 2045 TATCTCCAAATCAACACTAAGTCTGTTGAGAGATTTATGATGCTTTCGTAAGG 2104
QY 1801 CCTTGTGAGGCTTACCAACACTAGACGATCTTCAATAATAAAGA 1852
Db 2105 CCTTGTGAGGCTTACCAACACTAGACGATCTTCAATAATAAAGA 2156
```

```
RESULT 3
AE001654/LOCUS 10421 bp DNA linear BCT 01-DEC-2000
DEFINITION Chlamydia pneumoniae section 70 of 103 of the complete genome.
ACCESSION AE001654 AE001363
VERSION AE001654.1 GI:4377031
KEYWORDS
SOURCE Chlamydia pneumoniae CWL029
ORGANISM Chlamydia pneumoniae CWL029
REFERENCE 1 (bases 1 to 10421)
AUTHORS Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R. W.,
Olinger, L., Grimwood, J., Davis, R. W. and Stephens, R. S.
TITLE Comparative genomes of Chlamydia pneumoniae and C. trachomatis
JOURNAL Nat. Genet. 21 (4), 385-389 (1999)
MEDLINE 99206606
PUBMED 10192388
REFERENCE 2 (bases 1 to 10421)
AUTHORS Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Olinger, L.,
Grimwood, J., Davis, R. W. and Stephens, R. S.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1998) Program in Infectious Diseases, University
of California, 235 Earl Warren Hall, Berkeley, CA 94720, USA
FEATURES
location/Qualifiers
1..10421
/organism="Chlamydia pneumoniae CWL029"
/mol_type="genomic DNA"
/strain="CWL029"
/db_xref="taxon:115713"
133..2751
/gene="CPn0727"
133..2751
/gene="CPn0727"
/codon_start=1
/transl_table=1
/product="C7619 hypothetical protein"
/protein_id="AAD18866.1"
/db_xref="GI:4377032"
/translation="NSTFSIQNRLRTISGESTRIKLDHKYGFDPSPVAINLEELN
```

```
SGIYALRLHMLQSENTNVALLNPNTIIPPTSTWTDYKHSRPOASSRAPSSOPT
DIVSAAALVLDGLLVASTEIDLGAISTISTVROLMASVLGTLTAEOK
VFFSYVSEPNLLEHVKQRAAEIOAKQBEIKAVLEAKGVSTREIEAILKEPFIY
AADFFKEFTEEPHLYRAKVGAPIQEMNENAIQLPPTPAITPDNVNVMGMLSTII
LOADIVLQAPALGQDQIITLQPLVPLVDKTTFKAEFDLIYATQPLNTASLKL
LYTDROIAGYRGKITKVYQNSIQLSETKRVVNNRSMLETSQMFQACQNCFTWIS
QANALNIALTNKYISAVLITSMYGGLLCLSYMYERLADDEKAIKDKSVNYLPIHI
VVGSGWNGWIAKMAAYQELAYSLGTASODQIKAYLQTRGNEFKATRHPFNHIGD
QMYFANETVFGNCLTANGAIQPDUGGIFIREMTNVGVEADYVNAQRILNEFWTA
ATAHVLQQLQAEIOKQADDIDPGKASFTENKRFKFAVAAMITSESLDALISMLNSQ
LPQEAFLKPLIEEINFNALANALNSLIQTNEFTSTVYSSLSYLVSQKGTGNLF
AGDYSTLLAARERREYIYRDTRCKQALNVLNGLKINSPLPGATSAQCKMLNATT
VYQYSVLNLTNLVLESLLAGLKMLOTTSNNKYDKSVFKIESPDWPTLALLESF
LISGFENISATGGLGPLFTQVQSDQQTYSQGTQQLNLQNQMNTTIQEWTLVSTSMQ
VLNGLISLAGAIYSN"
complement (3078..5033)
/gene="CPn0728"
complement (3078..5033)
/gene="CPn0728"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_1 (CT622)"
/protein_id="AAD18867.1"
/db_xref="GI:4377033"
/translations="MVNPIGPPIDETERPPADLSAQGLEASAANKSAQRIAGAE
APKESKTSQSVBWSILRSVAVNALMSLADKLGIASSNSSTSRSDVDVDTTATP
PPTFDYKTAQATYDTITSTSLADIAALVLODAVTNIKDTAETDETAIAEAW
ETKNAADVKGAGQITELAKTASNDQALDLSGLTKLTSPDLLQALLQSVANNKAAELL
KENDPVEFGKTPAIAQSLVDQATQIENDQNAIRDAYFAGNQAGAVENAKSN
NSISNDSAKAAIATAKQIAEAKKFPDPIQEAPOMWIOAKDLKNTKPDQSDV
PNPQTVGGSKQCGSSIGSRVSMILLDDAENETASILMSGFQIMHNTENPDQSA
QOELAAQAAKAGDDSAALADAKALEAALGKAGQOQILNALQASAAVGA
GVPAASSTIGSSVKOLYKTSKSTQKQISAGYDAYKSIINDAVQANDATRDVI
NVSTPALTSVRPAREARGPKTQALARVIGNSRTIGDYVSQALQSVMLIQ
SNPQANNEIRKLTSLAKPPQPGIPYQVLSNDSTQKFIKLESLEPAESRTAAEIK
ALSFTNSLFIQVVLVNIYSYGLQ"
complement (5145..6413)
/gene="CPn0729"
complement (5145..6413)
/gene="CPn0729"
/codon_start=1
/transl_table=11
/product="CHLPN 76 kDa Homolog_2 (CT623)"
/protein_id="AAD18868.1"
/db_xref="GI:4377034"
/translations="MKKQVQWLASVLLALTISGYAELPLSEKVKSHYTTLDDEVK
DYLSKGFVETRKQDGVLRAGDVRAWRWLFREDIKNPSDKDKTNPLVNRYSSEYL
YIDYRAERNWLSKMMWTAAGENTAAGVDINRAFLGYFYKPNPRTDFFMIEGRS
GLGLPFESEVQFSQNDGLHIYMTRELSKDYQVIVHGGPFPVNNMTKKHYAVVEGI
LNRLPKQFFVKCSVVDWNTFVPSSTTEKAATNAMYKYCVQWMLVGRKHSQVWING
QKPLTYGAFPLMNLAKATKTLNGKENLAWFTGTLGLRKGADWSATVRYVEVA
LSVPEIDVSGIGRNLLKFWFAQAIANVDPKEANGFTNYKGFSAVWYGTITDSLSP
AYGAYSXANDKXGSDFTFRKDLGIISAF"
complement (6516..8159)
/gene="mvn"
/notes="synonym: CPn0730"
complement (6516..8159)
/gene="mvn"
/codon_start=1
/transl_table=11
/product="Integral Membrane Protein"
/protein_id="AAD18869.1"
/db_xref="GI:4377035"
/translations="MSRKDNVSLARSIENILSGTFCSRITGIFREIMATYFGADPI
VAAPWLGRTVFFLRKILGGLILEQAFIPHEFLRAQSLBRAAPFFRFRSLIKGSLI
IFTLLEAVLWVLQYBEGTYDMILLTPCGIFIMNNGALLHCENKFFGVG
LAPVWNIIWIFFYAAHSDPRERIIGLSVALVIGFFELITVPVWMLKELBAKSP
POEHSVRALLAPLSGLITTSIFQLNLLSDICLARYVHEITGLYLSKIYQLYPI
LFGFGVTVLLPAISRCVQREDHERGLKMKFVLTLTMSVMIIMTAGLILALPGVRV
LYEHGLPQSAYVAIVRVLYSGYASIIIPMALAPLVSLFVAQROYAVPLTIGIGTALA
NIVLSVLGRWLVDVSGISATISATVQVLYFLWYSSKRLPMYKLLWBSIRRSIK
VMGTLMACMITLGLNLTQTYVIFLNPPLTFLAWPLISSITAAQAFIASESCIFLAFI
FGFAKLARVEDLINLASFEYWRGQRLQROHVMQDTQN"
```

gene	8143..8319	Db	4469	GACTCTTTAGGTAACTGACTTCTTTCGACCTCTTACAGGCTGCTCTTCTCCCAATCTGTA	4410
CDS	/gene="CPn0731"	QY	421	GCAACAATAACAAAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	480
	8143..8319	Db	4409	GCAACAATAACAAAGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA	4350
	/codon_start=1	QY	481	GGGAAACCGCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTACACGACACAG	540
	/product="hypothetical protein"	Db	4349	GGGAAACCGCTGCAATTCCTCAATCTTTAGTTGATCAGACAGATGCTACACGACACAG	4290
	/protein_id="AADI8870.1"	QY	541	ATAGAGAAAGATGGAAATGCGATTAGGATGCATATTTTGACGACAGAACCTAGTGA	600
gene	8654..9535	Db	4289	ATAGAGAAAGATGGAAATGCGATTAGGATGCATATTTTGACGACAGAACCTAGTGA	4230
CDS	/note="synonym: CPn0732"	QY	601	GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACAGTAAACAGCA	660
	8654..9535	Db	4229	GCTGTAGAAATGCTAAATCTAATAACAGTATAAGCAACATAGATTACAGTAAACAGCA	4170
	/gene="nfo"	QY	661	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCCAAT	720
	/codon_start=1	Db	4169	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGACTCTCCAAT	4110
	/product="Endonuclease IV"	QY	721	CTTCAAGAGCGGAACAAATGCTAATACAGGCTGACAAAGATCTTAAAAATATCAAACT	780
	/protein_id="AADI8871.1"	Db	4109	CTTCAAGAGCGGAACAAATGCTAATACAGGCTGACAAAGATCTTAAAAATATCAAACT	4050
	/db_xref="GI:4377037"	QY	781	GCAGATGCTTCTGATGTTCCCAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACAGGA	840
gene	complement(9654..10283)	Db	4049	GCAGATGCTTCTGATGTTCCCAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACAGGA	3990
CDS	/gene="rs4"	QY	841	AGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTAGATGATGCTGATAAATGAGACCGCT	900
	/note="synonym: CPn0733"	Db	3989	AGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTAGATGATGCTGATAAATGAGACCGCT	3930
	complement(9654..10283)	QY	901	TCCATTTTGTGATGCTGGGTTTCGTGATGATGATGATGATGATGATGATGATGATGATGAT	960
	/gene="rs4"	Db	3929	TCCATTTTGTGATGCTGGGTTTCGTGATGATGATGATGATGATGATGATGATGATGATGAT	3870
	/codon_start=1	QY	961	TCTCAAGCTGCCCAACAGGAGCTCCGACACAGCTAGACAGCGCAAGACCCGCTGAGAT	1020
	/transl_table=11	Db	3869	TCTCAAGCTGCCCAACAGGAGCTCCGACACAGCTAGACAGCGCAAGACCCGCTGAGAT	3810
	/product="S4 Ribosomal Protein"	QY	1021	GCAGTGTCTGCTGACGCTGCGATGCTCAGAAAGCTTTTGAAGCGCTCTAGTAA	1080
	/protein_id="AADI8872.1"	Db	3809	GCAGTGTCTGCTGACGCTGCGATGCTCAGAAAGCTTTTGAAGCGCTCTAGTAA	3750
	/db_xref="GI:4377038"	QY	1081	GCTGGGCAACACAGGGGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTG	1140
	/translation="MARYCGPKNRVARRRGANIFGRSNPLKKPHPPGQHWORKKK	Db	3749	GCTGGGCAACACAGGGGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTG	3690
	SDYGLQLEKQKAKYAGIMMERQKLVKAEVHKGNVAQMFLEFECRLDNMYRM	QY	1141	AGGCGAGGAGTTCTCCCGCTGCGACAGCTTCTATAGGCTCATCTGTAAACAGCTTTAC	1200
	GKATTFIAAQQLVAHQHLLVNGRVDRFFLRPGMQISLKESKRLQSVKDALESKD	Db	3689	AGGCGAGGAGTTCTCCCGCTGCGACAGCTTCTATAGGCTCATCTGTAAACAGCTTTAC	3630
	ESSLPYSYISLDKTFKXGELLVSPQDQIEAQLPLINISVVCFFLSHRT"	QY	1201	AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTATGATGCT	1260
	ORIGIN	Db	3629	AAGACCTCAAAATCTACAGGTTCTGATTATAAACAACAGATATCAGAGGTTATGATGCT	3570
	Query Match	QY	1261	TACAAATTCATCATGATGCTTATGTAGGCGACGAAATGATGCGACTCGTGTATGATA	1320
	Best local similarity	Db	3569	TACAAATTCATCATGATGCTTATGTAGGCGACGAAATGATGCGACTCGTGTATGATA	3510
	Matches 1850; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	QY	1321	AACAAATGTAAGTACCCCGCTCTCACAGATCGGTTCCCTAGAGCACGAAACAGAGCTCGA	1380
	1	Db	3509	AACAAATGTAAGTACCCCGCTCTCACAGATCGGTTCCCTAGAGCACGAAACAGAGCTCGA	3450
	4829	QY	1381	GGACCAAGAAAAACAGATCAAGCCCTCGCTAGGAGTATTTCTGGCAATAGAGAACTCTT	1440
	61	Db	3449	GGACCAAGAAAAACAGATCAAGCCCTCGCTAGGAGTATTTCTGGCAATAGAGAACTCTT	3390
	4769	QY	1441	GGAGATGCTTATAGTCAAGTTTCGGGCACTACAAATCTGTAATGAGATCACTCAGTGAAT	1500
	121	Db	3389	GGAGATGCTTATAGTCAAGTTTCGGGCACTACAAATCTGTAATGAGATCACTCAGTGAAT	3330
	4709	QY			
	181	Db			
	4649	QY			
	241	Db			
	4589	QY			
	301	Db			
	4529	QY			
	361	Db			

```
1501 CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTTCCA 1560
Db CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTTCCA 3270
1561 CAGTTTGGCTATCCTTATGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1620
Db CAGTTTGGCTATCCTTATGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 3210
1621 TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGTGAATATAAGACACATTTCTTT 1680
Db TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGTGAATATAAGACACATTTCTTT 3150
1681 GAAACGAACTCCTTGTATTATTCAGCAGTGCTGCTCAATATCGGCTCTCTATATCTGGT 1740
Db GAAACGAACTCCTTGTATTATTCAGCAGTGCTGCTCAATATCGGCTCTCTATATCTGGT 3090
1741 TATCTCAATAACACACACCTTAAGTGTGCTTGGAGAGATTTATATGCTTGTGTAAGG 1800
Db TATCTCAATAACACACACCTTAAGTGTGCTTGGAGAGATTTATATGCTTGTGTAAGG 3030
1801 CCTTTGTTGAGGCTTACCACACACTAGACGATCTTCAATAATAATAAAGA 1852
Db CCTTTGTTGAGGCTTACCACACACTAGACGATCTTCAATAATAATAAAGA 2978

RESULT 4
AE002165 12548 bp DNA linear BCT 30-MAY-2000
LOCUS Chlamydomophila pneumoniae AR39, section 2 of 94 of the complete
DEFINITION genome.
ACCESSION AE002165
VERSION AE002165.1 GI:7188948
KEYWORDS Chlamydomophila pneumoniae AR39
SOURCE Chlamydomophila pneumoniae AR39
ORGANISM Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
REFERENCE 1 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gail, S.R., Heidelberg, J.F.,
White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T.,
Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B.,
Bowman, C., Dodson, R., Gwinn, M., Nelson, W., DeBoy, R., Kolonay, J.,
McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
pneumoniae AR39
JOURNAL Nucleic Acids Res. 28 (6), 1397-1406 (2000)
MEDLINE 20150255
PUBMED 10684935
REFERENCE 2 (bases 1 to 12548)
AUTHORS Read, T.D., Brunham, R.C., Shen, C., Gail, S.R., Heidelberg, J.F.,
White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T.,
Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B.,
Bowman, C., Dodson, R., Gwinn, M., Nelson, W., DeBoy, R., Kolonay, J.,
McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.
Direct Submission
JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
MEDLINE Medical Center Dr, Rockville, MD 20850, USA
PUBMED

FEATURES
source
1..12548
/organism="Chlamydomophila pneumoniae AR39"
/mol_type="genomic DNA"
/strain="AR39"
/db_xref="taxon:115711"
/notes="synonym: Chlamydia pneumoniae AR39"
297..965
/gene="CP0011"
297..965
/gene="CP0011"
/notes="similar to SP:P31218 GB:X71492 PID:296947 GB:U00096
PID:I736770; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="uridine kinase"

gene
CDS
1501 CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTTCCA 1560
Db CCTCAAGCGAATATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTTCCA 3270
1561 CAGTTTGGCTATCCTTATGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 1620
Db CAGTTTGGCTATCCTTATGCACTTTCTAATGACTCTACACAGAGTTTCATAGCTAAA 3210
1621 TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGTGAATATAAGACACATTTCTTT 1680
Db TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGAGTGAATATAAGACACATTTCTTT 3150
1681 GAAACGAACTCCTTGTATTATTCAGCAGTGCTGCTCAATATCGGCTCTCTATATCTGGT 1740
Db GAAACGAACTCCTTGTATTATTCAGCAGTGCTGCTCAATATCGGCTCTCTATATCTGGT 3090
1741 TATCTCAATAACACACACCTTAAGTGTGCTTGGAGAGATTTATATGCTTGTGTAAGG 1800
Db TATCTCAATAACACACACCTTAAGTGTGCTTGGAGAGATTTATATGCTTGTGTAAGG 3030
1801 CCTTTGTTGAGGCTTACCACACACTAGACGATCTTCAATAATAATAAAGA 1852
Db CCTTTGTTGAGGCTTACCACACACTAGACGATCTTCAATAATAATAAAGA 2978

/protein_id="AAF37907.1"
/db_xref="GI:7188950"
/translation="MLMLMLIIGITGGGAGKTLTQNIKEIFGEDSVICODNYK
DRSHYTPERANLIMDPADFDLILSDIKELKNEIYCAFPVFLVGRSKTELET
IYPSVILVEGILVFENQELMDIRIIVDTADERILRRMVRDVBQGSVDICIMS
RYLSVKPHEKFIETPRKLDIIVHGNVRQNVNLSQIKNHLNALESDETYYM
VNSK"
complement(1047..2021)
/gene="CP0012"
complement(1047..2021)
/gene="CP0012"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37908.1"
/db_xref="GI:7188951"
/translation="MEKYYALAYVYITVDNPHEIALHKFLIEDVSCRTYISEQ
GINGQSGVEPHAEALYMOALKSRPNESKIKELHHIKENIFPRITVKYKELAALGCE
VLSKQAKHISFOEWEHEKLEQNRCLILDVNNYEWKIGHFDNATLPIQOTFRFPBVA
EKLAQCDPETTPVMYCTGGIRCELYSVLLEKPKFVQLDGGVIAFGQQVGTGKW
LGKLFVDDRLAIPIDEDPDVAPIAECCHCOTPSDAYNCANTDCNALFLCDECIH
QHGCGCECSQSPRVKFDSSRGKPFRAHLCEISENSEASCCLI"
2238..2867
/gene="CP0013"
2238..2867
/gene="CP0013"
/notes="similar to SP:P81288; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="ribosomal protein S4"
/protein_id="AAF37909.1"
/db_xref="GI:7188952"
/translation="MARYCGPKNRVARFQANIFGRSNPLLLKKHPHGGHGMQRKKK
SDYQLOLEEKQKLKACYGIMKQLVAKFVKHKGQNVAFELRFECDLNNVYRM
GPAKTPFAAQQLVAHGHIILVRRVRDRRFFLRPGMQISLKEKSKELQSVKDALESKD
ESSLPYSYISLDKGTGKGEILLVSPQDQIEAQLPLPINISVCFELSHRT"
complement(2986..3867)
/gene="CP0014"
complement(2986..3867)
/gene="CP0014"
/notes="similar to SP:P12638 GB:M22591 PID:146954
PID:405898 GB:U00096; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="endonuclease IV"
/protein_id="AAF37910.1"
/db_xref="GI:7188953"
/translation="MKVLPPSPILGAGTSTAGGLKNAIYEGRDIGAGTVOIFTAQC
RQWRALKEVIEDPKAALKETDLSYIMSHAGYLINPGAPDPVILEKSEIGIYOETL
DCITLGSFVNPFHFGAALKSKEDCMNKIVSSFSAPLFDSSPPLVLVLETTAGQST
LIGSNFELGVLQNLKQIPICVDTCHIFAGYDITSPQGWEDVINEFDEIVGLS
YLRAFINDSMFPLGANKDRHAPGLGEGYIKGSFKFLMTDRTRKPKYLETPGPPEN
WQKEICELLKFSKNRDS"
complement(4202..4378)
/gene="CP0015"
complement(4202..4378)
/gene="CP0015"
/notes="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF37911.1"
/db_xref="GI:7188954"
/translation="MSFLLIKRSPPLKQDYLFDITSPYTSSTITIGGSYFKLNKASL
QSSTLRSLRSIIIS"
4362..6005
/gene="CP0016"
4362..6005
```



```

/gene="CP0016"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37912.1"
/db_xref="GI:7188955"
/translation="MSRKDNVSLARSIFENILSGTFCRSIRIGIFREIAMAIFYGADPI
VAFWLGRVTVFFKILGVLLEQAFIPHEFLRQSLDRAAFRRFRSLIKGSTI
IPTLLIEVWLVVLYVEEGYDMILLTMIILPCGIFLMNVNNGALLHCENKFFGV
LAPVVNIIWFFVIAARHSDPRERIIIGLSVALVIGFFEMLIIVPGVVKFLEAKSP
POEHDVALLAPLSGLITSSIFOLNLLSDICLARYVREIGPLVLMYSLKIYQPIH
LEGQVFFVLLPAISRCVOREHERGLKMKFVLTLTMSVMIITAGLLILLAPGVRY
LIEHGLFQSQSAVIVRLRGASIIIPMALAPLVSVLFPYARQRYAVPLFVIGITALA
NIVLSVLGRWVLLKDVSGISYATSIWVQVLYLMYISKRPLMTSKLWESIRRSIK
VMGTMLACMIITGLNIIQTQTVYIFLNPFLTFLAWPLSSITTAQAIATLSESCIFLAFI
FGFAXLLRVEDUINLASFEYMRGQRLQQRHVMDQTN"
6108..7376
/gene="CP0017"
6108..7376
/gene="CP0017"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37913.1"
/db_xref="GI:7188956"
/translation="MKQYVQWLASVLLALTISSYAEPLFSEQVKSHYTTTLDEVK
DYLGRGFVETKQGVILRIGDVRARWLYFREDIKNPSDKDKNPLRPNRYRSEFYI
YIDYRAERWESKMNWTAIAGDENTAAAGVDINRAFLVYRKNPDRDFFMEIGRS
GIGDLFESVQSQSNFDGLHYWTRELSKDPYQVTVHGGPFVWNTKHYAVWVEGI
LAKPLQFVKSVDWNTFVSESTTEKATNANKYKICVQWHLVKGHSQVPLNG
QKKPLYLYGAFILNPLAKATKTLNGLNWLFIPTGLGLRRAGDWSAIVRYEYEA
LSVPEIDVSGIGRNLKWFQAIAANYDPPKEANGFTNYKGFSAIYMYGITDSLSFR
AYGAVSKPANDKLGSDFTFRKFDLGIISAF"
7488..9443
/gene="CP0018"
7488..9443
/gene="CP0018"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37914.1"
/db_xref="GI:7188957"
/translations="MVPNIPGPGPIDETERTPPADLSAQGLEASANKSAEQAIRIAGE
AKPKESKTSVERWSILRGAVNALMSLADKLGIASSNSSTSRSDVDSTTATATP
PPTFDYKTOACTAYDTLTFTSLADIAALVLODAVNTIKDTAATDEBTAIAAEW
ETKNADAVKVAQITELAKYASDNQALISLGLTSFDLLQALLOSANNKKAELL
KEMQDNVPVPGKTPAQLSLVDQTDATQIEKDGNAIRDPAFAGQNASGAVENAKSN
NSISNDTSAKAALATATQIAEAKKFPDPSPILOEAEQWVIQAEKDLKNIKADGSDV
PNPPTVGGSKQGSIGSIRVSMILLDAENETASIMSGFRQIMHNPFPDSQAA
QOELAAQAAGAAGDAAAALDAKALEAALGRAGQOQGLNALGQITASAAYVSA
GVPPAASSTGSSVKOLYKTSKSTGDYKTOISAGDYAKSINDYGRARNDAITROVI
NNVSTALTRSVPKRTAEARKPEKTDQALARVISGNSRTLGDVYSQVLSALQWQIIQ
SNPQANNEIRKQTSVATKPPQFGYPVOLSNDSTQKFKAKLESFLAESRTPAABIK
ALSFETNSLFIQQLVNLIGLSYGIQ"
complement(9770..12388)
/gene="CP0019"
complement(9770..12388)
/gene="CP0019"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF37915.1"
/db_xref="GI:7188958"
/translation="MSTFSIQNRLRTISGESIRIILDKHYSGFDRSPVPAINLEELN
SGIYALRHLMNALQSENTVAALLNPNNTIFPTTSWTDYKHSRQASSPRAPSSQTP

```

```

DIVSAAALALVLVIDGGIAELVASVTEIDLGALSTISTVRQLMASYLGLTTLTAQEQK
VVFSSVPSKNELEHVKQKAAEIOAKOBEIKAVLEAKGVSTEEHAILKVPDIY
AADPFKSEFIEPHLHTYPAKGAPIQEMNENAIQLLPTPAPTNNVNEVNGMNTLSTI
LOADDDAIKOPALGGQBEIITILQTLVPLVDKTTFFKABEDLLYTATQLPNTASLKL
YLTDRQIAEYRKITKVYQNSIQNLSERKVRVNNRSMLETQKLSMFQQAQCNCFWTWIS
QANALNTAITNKYISAVLTTSMEYMGGLLSUSMYERLADDEKKAIFDKSVNEYLPFIH
VVGGSWNGWIAKMAAYQELAYSLGTAVTSDQIKAYLQTRGNEFKATREHFFHNGD
QMYQFANETVFNCLTTANGAIOPDLGGFIREAMTVNGTVEADYVSNQARILNBNFTA
ATAHVQLQLQIAELQKADDLPKASFTENRKXPAVAAMITSESLGALLI.SMIINSQ

Query Match      99.8%; Score 1848.8; DB 1; Length 12548;
Best Local Similarity 99.9%; Mismatches 0; Gaps 0;
Matches 1850; Conservative 0; Prid. No. 0;

QY      1  ATGAGTCGTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGCTCTTCTTACTAGCAGA 60
DB      1  ATGAGTCGTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGCTCTTCTTACTAGCAGA 7751
QY      61  TGTGCAGACGTGGACTCAACGACGACGACCGACCTACGCTCTCTCCACCCACGCTTTGAT 120
DB      61  TGTGCAGACGTGGACTCAACGACGACGACCGACCTACGCTCTCTCCACCCACGCTTTGAT 7811
QY     121  GATTATAGACTCAAGCGCAACAGCTTACGATATCTATTTTACCTCAACATCAGTACGT 180
DB     121  GATTATAGACTCAAGCGCAACAGCTTACGATATCTATTTTACCTCAACATCAGTACGT 7871
QY     181  GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATATAAGGATACAGCG 240
DB     181  GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACTAATATAAGGATACAGCG 7931
QY     241  GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGAGTT 300
DB     241  GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATGAGTT 7991
QY     301  AAAGTTGGCGCCAAATACAGAAATAGCGAAATATGCTTCGGGATAACCAACGATTCCT 360
DB     301  AAAGTTGGCGCCAAATACAGAAATAGCGAAATATGCTTCGGGATAACCAACGATTCCT 8051
QY     361  GACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTGTA 420
DB     361  GACTCTTTAGGTAAACTGACTTCCTTCGACCTCTTACAGGCTGCTCTTCTCCAACTGTA 8111
QY     421  GCAACAATTAACAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 480
DB     421  GCAACAATTAACAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA 8171
QY     481  GGGAAAACCCCTGCAATGCTCAATCTTGTAGTTGATCAGACAGATGCTCAGCGACACAG 540
DB     481  GGGAAAACCCCTGCAATGCTCAATCTTGTAGTTGATCAGACAGATGCTCAGCGACACAG 8231
QY     541  ATAGAGAAAGATGGAAATCGGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGA 600
DB     541  ATAGAGAAAGATGGAAATCGGATTAGGGATGCATATTTTGCAGGACAGAACGCTAGTGA 8291
QY     601  GCTGTAGAAAAATGCTAAATCTTAATACAGTATAAGCAATAGATTACAGTAAAGCAGA 660
DB     601  GCTGTAGAAAAATGCTAAATCTTAATACAGTATAAGCAATAGATTACAGTAAAGCAGA 8351
QY     661  ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTCCCGCACCTCTCCAAIT 720
DB     661  ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTCCCGCACCTCTCCAAIT 8411
QY     721  CTTCAAGAGCGGAACAAATGGTAATACAGGCTGAGAAAAAGATCTTAAAAAATATCAACCT 780
DB     721  CTTCAAGAGCGGAACAAATGGTAATACAGGCTGAGAAAAAGATCTTAAAAAATATCAACCT 8471
QY     781  GCAGATGGTTCGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACAGGA 840
DB     781  GCAGATGGTTCGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACAGGA 8531
QY     841  AGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTAGTAGTATGCTGCTGCTGCTGCTGCT 900
DB     841  AGTAGTATTGGTAGTATTGCTGTTCCATGCTGTTAGTAGTATGCTGCTGCTGCTGCTGCT 8591

```


KEYWORDS		Chlamydia pneumoniae TW-183
SOURCE		Chlamydia pneumoniae TW-183
ORGANISM		Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia
REFERENCE		1 (bases 1 to 300066)
AUTHORS		Geng, M.M., Schubmacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T., Essig, A., Marre, R. and Melchers, K.
TITLE		The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis
JOURNAL		Unpublished
REFERENCE		2 (bases 1 to 300066)
AUTHORS		Geng, M.M., Schubmacher, A., Muehldorfer, I., Bensch, K.W., Schaefer, K.P., Schneider, S., Pohl, T. and Melchers, K.
TITLE		Direct Submission
JOURNAL		Submitted (28-MAY-2002) RDR/IT, RDR/FG, RDR/BT, ALTANA Pharma, Byk-Gulden-Str. 2, Konstanz 78467, Germany
FEATURES		Location/Qualifiers
source		1..300066
		/organism="Chlamydia pneumoniae TW-183"
		/mol_type="genomic DNA"
		/strain="TW-183"
		/db_xref="taxon:182082"
		/note="type strain"
		biovar: TWAR
		126..797
		/locus_tag="CpB0543"
		126..797
		/locus_tag="CpB0543"
		/note="similar to hypothetical protein CT433 - Chlamydia trachomatis"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="AAP98472.1"
		/db_xref="GI:33236384"
		/translation="MTLYGLNQKTKARKYQAHYLPILTFFYNAKSTQNKRALQFLPQ ATHVLTSPSSTHLRSMTSLSKATKTKYLCIGESTKERLLSLGQVKYVATQ EIAEGIFLLQALPSSARILYPHSSILARVIRFLYRNRTFFSYHYTVKRLKKNL LSYKXILLTSPSTVRAFAKIFRFEKTIWQCGRMTLQBFQKFSQKQVSLLETGLK SRSP"
		862..1194
		/locus_tag="CpB0544"
		862..1194
		/locus_tag="CpB0544"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="AAP98473.1"
		/db_xref="GI:33236385"
		/translation="MASSATPGDGTAPSLFPLATPRVNFKLALFVTTIALIUVIAL IATVIALGLCHPLCSFIFLTAIPLYISRYICSHVARNVYIALDVVPDHSKLQDMRS HSRIFSDR"
		1242..2321
		/locus_tag="CpB0545"
		1242..2321
		/locus_tag="CpB0545"
		/codon_start=1
		/transl_table=11
		/product="hypothetical protein"
		/protein_id="AAP98474.1"
		/db_xref="GI:33236386"
		/translation="MSGPSRTESSQVSLSYVPRDEIAPKQFTIAKISTLAILASL ALGALVAGISLTIVLGNPVFLALLITTFALSVVTFVYHMTSKVSNWOKVLEQNFK PLKAWQKXNDVCSNEMQFNNHLPKFKVAICTDASQFQPTFLTGLVIEQNQST GIIFNPVGNLNDTAINSLSTLYLKDKSVNDTCQREBPAGKEDPSPTEVIV KALESHAAVALPFTSVYVFPPEILPEKGTGYWNTQTAQFCKRALLDAIQNTAUR YQPSRLVILQDPENTIESQSR8EE"
		2661..3425
		/locus_tag="CpB0546"
		2661..3425
		/locus_tag="CpB0546"

gene

CDS

gene

CDS

gene

CDS

gene

CDS

QY	901	TCCATTTTGATGCTCGGTTTCCTCAGATGATTCACATGTTCAATACCGAAATCTCGAT	960
DB	8592	TCCATTTTGATGCTCGGTTTCCTCAGATGATTCACATGTTCAATACCGAAATCTCGAT	8651
QY	961	TCTCAAGCTGCCAACAGGAGCTCGCAGCACAAGCTAGACGCGGCTCGAGAT	1020
DB	8652	TCTCAAGCTGCCAACAGGAGCTCGCAGCACAAGCTAGACGCGGCTCGAGAT	8711
QY	1021	GACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGTAA	1080
DB	8712	GACAGTGTCTGCTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGTAA	8771
QY	1081	GCTGGGCAACAAAGGCACTACATGCTTTAGGACAGATCGTCTGCTGCTGTG	1140
DB	8772	GCTGGGCAACAAAGGCACTACATGCTTTAGGACAGATCGTCTGCTGCTGTG	8831
QY	1141	AGCGCAGAGTTCCTCCGCTGCAGCAAGTTCCTATAGGTCATCTGTAAACAGCTTTAC	1200
DB	8832	AGCGCAGAGTTCCTCCGCTGCAGCAAGTTCCTATAGGTCATCTGTAAACAGCTTTAC	8891
QY	1201	AAGACCTCAAAATCTACAGGTTCTGATTAATAACACAGATATCAGCAGGTTATGATCT	1260
DB	8892	AAGACCTCAAAATCTACAGGTTCTGATTAATAACACAGATATCAGCAGGTTATGATCT	8951
QY	1261	TACAAATCCATGATGCTATGTTAGGACGCAAGTTCATGAGGTCATCTGTAAACAGCTTTAC	1320
DB	8952	TACAAATCCATGATGCTATGTTAGGACGCAAGTTCATGAGGTCATCTGTAAACAGCTTTAC	9011
QY	1321	AACAATGTAGTACCCCGCTCTCACAGATCCGTTCTAGAGCAGCAAGAGCTCGA	1380
DB	9012	AACAATGTAGTACCCCGCTCTCACAGATCCGTTCTAGAGCAGCAAGAGCTCGA	9071
QY	1381	GGACCAAGAAAAACAGATCAAGCCCTCGTAGGTCGATTTCTGCAATAGCAGAACCTCTT	1440
DB	9072	GGACCAAGAAAAACAGATCAAGCCCTCGTAGGTCGATTTCTGCAATAGCAGAACCTCTT	9131
QY	1441	GGAGATGCTATAGTCAAGTTTCGGCTCAATCTGTAAGTCAGATCACTCAGTCGAAT	1500
DB	9132	GGAGATGCTATAGTCAAGTTTCGGCTCAATCTGTAAGTCAGATCACTCAGTCGAAT	9191
QY	1501	CCTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTTCCA	1560
DB	9192	CCTCAAGCGAATAATAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTTCCA	9251
QY	1561	CAGTTTGGCTATCTTATGTGCACTTTCTAATGACTCTACACAGAGTTTATAGCTAAA	1620
DB	9252	CAGTTTGGCTATCTTATGTGCACTTTCTAATGACTCTACACAGAGTTTATAGCTAAA	9311
QY	1621	TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATATAAGAGCACTTCCCTTT	1680
DB	9312	TTAGAAAGTTTGTGCTGAAGGATCTAGGACAGCAGCTGAAATATAAGAGCACTTCCCTTT	9371
QY	1681	GAAACGAACCTCTGTTTATTACAGCAGTGTGCTCAATATCGGCTCTCTATATCTCGT	1740
DB	9372	GAAACGAACCTCTGTTTATTACAGCAGTGTGCTCAATATCGGCTCTCTATATCTCGT	9431
QY	1741	TATCTCCAAATACACACTAAGTTCGTTGGAGAGATTAATATGCTTCTGTAAGG	1800
DB	9432	TATCTCCAAATACACACTAAGTTCGTTGGAGAGATTAATATGCTTCTGTAAGG	9491
QY	1801	CGTTTGTGAGGCTTACCACACACTAGACGATCTTCAATATAAATAAAGA	1852
DB	9492	CGTTTGTGAGGCTTACCACACACTAGACGATCTTCAATATAAATAAAGA	9543

gene

CDS

gene

CDS

gene

CDS

gene

CDS

RESULT 5

AE017159/c

LOCUS

DEFINITION

Chlamydia pneumoniae TW-183, section 3 of 4 of the complete genome.

ACCESSION

AE017159

VERSION

AE017159.1

GI:33236383

300066 bp

DNA

linear

BCT 25-JUL-2003

QY	301	AAAGTTGGCGCGAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT	360	1381	GGACCAGAAAAACAGATCAAGCCCTCGTAGGTGATTTCTGGCAATAGCAGACTCTT	1440
Db	214097	AAAGTTGGCGCGAAATACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCGATTCTT	214038	213017	GGACCAGAAAAACAGATCAAGCCCTCGTAGGTGATTTCTGGCAATAGCAGACTCTT	212958
QY	361	GACTCTTTAGGTAAATGTACTTCCTTCGACCTCTTAAGGATGATGATTAACCAAGCGATTCTT	420	1441	GGAGATGCTTATAGTCAAGTTTTCGGCACTTACAAATCTGTAATCGAGATCACTCAGTCGAAT	1500
Db	214037	GACTCTTTAGGTAAATGTACTTCCTTCGACCTCTTAAGGATGATGATTAACCAAGCGATTCTT	213978	212957	GGAGATGCTTATAGTCAAGTTTTCGGCACTTACAAATCTGTAATCGAGATCACTCAGTCGAAT	212898
QY	421	GCAAAACAATAAAGCAGCTGAGCTTCTTAAGAGATGCAAGATTAACCAAGCGATTCTT	480	1501	CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGCGAGTGCAAAAGCTTCA	1560
Db	213977	GCAAAACAATAAAGCAGCTGAGCTTCTTAAGAGATGCAAGATTAACCAAGCGATTCTT	213918	212897	CCTCAAGCGAATAATGAGGAGATCAGACAAAAGCTTACATCGCGAGTGCAAAAGCTTCA	212838
QY	481	GGGAAACGCGCTCAATGCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACAG	540	1561	CAGTTTGGCTATCTTATGTCGAACTTCTTAATGATCTTACACAGAAGTTATAGTAA	1620
Db	213917	GGGAAACGCGCTCAATGCTCAATCTTTAGTTGATCAGACAGATGCTCAGCGACAG	213858	212837	CAGTTTGGCTATCTTATGTCGAACTTCTTAATGATCTTACACAGAAGTTATAGTAA	212778
QY	541	ATAGAGAAAGATGGAATGCGATTAGGATGCAATTTTCAGGACAGACGCTAGTGA	600	1621	TTAGAAAAGTTTCTTGTCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTCTCTT	1680
Db	213857	ATAGAGAAAGATGGAATGCGATTAGGATGCAATTTTCAGGACAGACGCTAGTGA	213798	212777	TTAGAAAAGTTTCTTGTCTGAAGGATCTAGGACAGCAGCTGAAATAAAAGCACTTCTCTT	212718
QY	601	GCTGTAGAAATGCTAAATCTTAATAAGCAATAGATTAAGCAATAGATTAAGCAATAG	660	1681	GAAACGAACTCTTGTATTATTCAGCAGGCTGCTGGTCAATATCGGCTCTCTATTTCTGT	1740
Db	213797	GCTGTAGAAATGCTAAATCTTAATAAGCAATAGATTAAGCAATAGATTAAGCAATAG	213738	212717	GAAACGAACTCTTGTATTATTCAGCAGGCTGCTGGTCAATATCGGCTCTCTATTTCTGT	212658
QY	661	ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAAAT	720	1741	TATCTCCAATAACAACACCTTAAGTGTTCGTTGGAGAGATTATTATGCTTTGGTAAGG	1800
Db	213737	ATCGTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGACTCTCCAAAT	213678	212657	TATCTCCAATAACAACACCTTAAGTGTTCGTTGGAGAGATTATTATGCTTTGGTAAGG	212598
QY	721	CTTCAAGAGCGGAACAAATGTAATACAGCTGAGAAAGATCTTAAATAATCAAACT	780	1801	CCTTTGTTGAGGCTTACCACACACACTAGAACGATCTCAATAATAAAGA	1852
Db	213677	CTTCAAGAGCGGAACAAATGTAATACAGCTGAGAAAGATCTTAAATAATCAAACT	213618	212597	CCTTTGTTGAGGCTTACCACACACACTAGAACGATCTCAATAATAAAGA	212546
QY	781	GCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGGA	840	RESULT 6		
Db	213617	GCAGATGTTCTGATGTTCCAAATCCAGAACTACAGTTGGAGGCTCCAAAGCAACAGGA	213558	AP002547/c		
QY	841	AGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT	900	LOCUS	300550 bp DNA linear BCT 25-APR-2003	
Db	213557	AGTAGTATTGGTAGTATTCGTTTCCATGCTGTTAGATGATGCTGAAATAGACCGCT	213498	DEFINITION	Chlamydia pneumoniae J138 genomic DNA, complete sequence,	
QY	901	TCCATTTTGTGCTGCTGGTTTCGTCAGATGATTCACATGTTCAATAGGAAATCCCTGAT	960	ACCESSION	AP002547.2 GI:10176693	
Db	213497	TCCATTTTGTGCTGCTGGTTTCGTCAGATGATTCACATGTTCAATAGGAAATCCCTGAT	213438	VERSION	AP002547.2	
QY	961	TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACAGCGAAAGCGCTGGAGAT	1020	KEYWORDS	Chlamydia pneumoniae J138	
Db	213437	TCTCAAGCTGCCAACAGGAGCTCGCAGCAACAGCTAGACAGCGAAAGCGCTGGAGAT	213378	SOURCE	Chlamydia pneumoniae J138	
QY	1021	GACAGTGTGCTGAGCGCTGGAGATGCTCAGAAAGCTTTAGAGCGGCTTAGGTAAA	1080	ORGANISM	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.	
Db	213377	GACAGTGTGCTGAGCGCTGGAGATGCTCAGAAAGCTTTAGAGCGGCTTAGGTAAA	213318	REFERENCE	1 Shirai, M., Hirakawa, H., Ouchi, K., Tabuchi, M., Kishi, F., Kimoto, M.,	
QY	1081	GCTGGCAACAAAGGCGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTG	1140	AUTHORS	Takeuchi, A., Nishida, J., Shibata, K., Fujinaga, R., Yoneda, H.,	
Db	213317	GCTGGCAACAAAGGCGATCTCAATGCTTTAGACAGATCGCTTCTGCTGCTGTTGTG	213258	TITLE	Natsushima, H., Tanaka, C., Furukawa, S., Miura, K., Nakazawa, A.,	
QY	1141	AGCGCAGAGTTCTCCGCTGCGAGAGTTCTATAGGCTCATCTGTAACACAGCTTAC	1200	JOURNAL	Ishii, K., Shiba, T., Hattori, M., Kuhara, S. and Nakazawa, T.	
Db	213257	AGCGCAGAGTTCTCCGCTGCGAGAGTTCTATAGGCTCATCTGTAACACAGCTTAC	213198	MEDLINE	Comparison of outer membrane protein genes omp and pmp in the whole	
QY	1201	AAGACCTCAAAATCTACAGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT	1260	PUBMED	genome sequences of Chlamydia pneumoniae isolates from Japan and	
Db	213197	AAGACCTCAAAATCTACAGTTCTGATTATAAACAACAGATATCAGCAGGTTATGATGCT	213138	REFERENCE	the United States	
QY	1261	TACAAATCCATCAATGATGCTTATGGTAGGCGACGAAATGATGGACTCGTGATGTGATA	1320	AUTHORS	J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)	
Db	213137	TACAAATCCATCAATGATGCTTATGGTAGGCGACGAAATGATGGACTCGTGATGTGATA	213078	TITLE	2 Shirai, M., Hirakawa, H., Kimoto, M., Tabuchi, M., Kishi, F., Ouchi, K.,	
QY	1321	AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGCTCGA	1380	JOURNAL	Shiba, T., Ishii, K., Hattori, M., Kuhara, S. and Nakazawa, T.	
Db	213077	AACATGTAAGTACCCCGCTCTCACAGATCCGTTCTTAGAGCAGCAACAGAGCTCGA	213018	REFERENCE	Comparison of whole genome sequences of Chlamydia pneumoniae J138	

```

FEATURES             Location/Qualifiers
     1..30050
     /organism="Chlamydomonada pneumoniae J138"
     /mol_type="genomic DNA"
     /strain="J138"
     /db_xref="taxon:138677"
     /note="synonym:Chlamydia pneumoniae (strain J138)"
     complement(21..794)
     /gene="dapF"
     complement(21..794)
     /gene="dapF"
     /codon_start=1
     /transl_table=11
     /product="diaminopimelate epimerase"
     /protein_id="BAA98725.1"
     /db_xref="GI:8978890"
     /translation="NAFYSPSTISKYFIYSGAGNRFLLGELTPEVEDVRFLOQETRV
GLYLYKPSQCAQILIFNSDGRSLRPMCMGNLRCALAHLSAKGSDISVSDGLYS
GYFYSDRLVDMTADVASVHRLESRPDLPEKVCVCHTGVPHAVILPEISLIDL
SILGPELVHOFTEPDSGVNVFVQLIGCOLRVRYERGVEGTAACTGALASALV
SNSYGWKESIHTWGGELMTVSQNRGVYLGQSVTRDL"
     complement(763..1338)
     /gene="clpP.1"
     complement(763..1338)
     /gene="clpP.1"
     /codon_start=1
     /transl_table=11
     /product="Clp protease"
     /protein_id="BAA98726.1"
     /db_xref="GI:8978891"
     /translation="MADGSRVHKLRIIEKELLEARRVFSEPVTEKSASDAIKLWYL
EKDQKPIVFVNSPGGSVDAGFAVMDIKMLTSPVTVTGLAASMSVLSLCAAP
GRFATPHSRIMHQPSIGSGPTGQATDLDIHAREILTKARLIIDVVEATNQPRDII
EKXADDMWTANEAQDFGLDGLIFSNL"
     complement(1358..2917)
     /gene="glyA"
     complement(1358..2917)
     /gene="glyA"
     /codon_start=1
     /transl_table=11
     /product="serine hydroxymethyltransferase"
     /protein_id="BAA98727.1"
     /db_xref="GI:8978892"
     /translation="MLKVPEKKPKFAIVEFTKVAVVSVLLHKFLFNASGKQSLAS
TAYLALDLHLNAPSIGRIIDELKSQSHLXMIASENYSSLSVGLANGNLLTKYC
EGPPKRFYSCCNDAIEWCVEETAKELPAACVQPHSGADANLLAVAILTHKV
QCPAVSKLGTNTLEETEEYTLKAEKSSVCLGSLNSGHLTHGNVRLNWSKLM
RCFPIYDNPTECFDAEISRLAKYKPKVLYAGSYSGRLNFAVLKQIAEDCSVL
WYDMAFAGLVAGGVFDEENPIPYADIVTTTHKTLRGRGLVLAETREYSLNKA
CPLMGSGPLPHVIAAKTVALKEALSDFDKYAHOVVNNARRLAERFLSHGLITGTL
DNHMVYDILGSLGSKIAEDILSSVGIARNRSLPSDAIGKWDTSGLRGTALTTL
GWGIDMEBEVADIIVKLNRLNLSCHVEGSSKKNKGELPEATAQEARDRVRLNLLRFP
LYPEIDLEALV"
     3107..3778
     /gene="CPj0522"
     3107..3778
     /gene="CPj0522"
     /codon_start=1
     /transl_table=11
     /product="CT433 hypothetical protein"
     /protein_id="BAA98728.1"
     /db_xref="GI:8978893"
     /translation="MTLYGLNOKTARKYQAHYLPILTLPFYAKSTPQNKARLQFLPQ
ATHVILTSPSSTHLFSLRMTSLSKATLTKTYLCTIGESTKERLLSFLQGVKVVATQ
ETAEAGIFFDLQALPFSRLIYPHSSLARPVIREFLYNRFTFFSYPHYTKPRKKNKI
LSKYKXIIITPSPTVRAFAKIFPRFPPEKTYCQGRMTLQEFQFSSQKQVSLLETIGK
SRTPSP"
     3843..4175
     /gene="CPj0523"
     3843..4175
     /gene="CPj0523"
     /codon_start=1
     /transl_table=11
     /product="hypothetical protein"
     /protein_id="BAA98729.1"
     /db_xref="GI:8978894"
     /translation="WASSATFGEDGTAPSLFPPTAPRPNFKLALFVTIATLWVIAL
IATTIATGLCHPLCSFILTALPIFYISRYCHYARNVYIALDVVPDHSKLODMRS
HSPIFSDR"
     4223..5302
     /gene="CPj0524"
     4223..5302
     /gene="CPj0524"
     /codon_start=1
     /transl_table=11
     /product="hypothetical protein"
     /protein_id="BAA98730.1"
     /db_xref="GI:8978895"
     /translation="MSGFSRTESQSVLSVPRKETAIPKKQFTIAKISTAILLASL
ALGALVAGISITVILGNFVFLALITLALFVTVTLVYHQTSKVSSNWQKVEQNFK
PIGKAWQGNVTCYNEMQFNHNLNPKFVAIQTASQFPQFTLTGLRVEKQST
GIIIPNVGPTNLIDNTATNLSTLYTLKDSKVMDTCKREGPAKGDPSFPTVRV
VKLPNEALDQTNLNLSSAEKSLILPTFLGHVCPKSELPNQKQVYRQALLAYENCL
KAAIESHAALVALPLFTSYVYVPPPEILPKGSTFYWDNQTAQACEKALLDALQNTALR
YQSRLLVILQDPFNITIEQSRSE"
     5642..6406
     /gene="CPj0525"
     5642..6406
     /gene="CPj0525"
     /codon_start=1
     /transl_table=11
     /product="CT398 hypothetical protein"
     /protein_id="BAA98731.1"
     /db_xref="GI:8978896"
     /translation="MHDALLSLTAQELDIKIMIRLMRVKKEHOKELAKVQSLKDIRR
KYQKELMENLKTQIRGENRIQELISQINKLENOQAAVKMDENALTOEMWTANK
ERSUEHOLSMDKQAGEDLIVLSKESLASTENSSSVIEKEIFESIKKINEBKAL
LEQRTLEKHAIPNELLSIYERLLNNKKORVVPIENRVCSGCHIILITPOHENVLRKDO
RLIFCEHCSRILYQWESQVNAQENSTAKRRRRRAAV"
     complement(6827..7816)
     /gene="kpsP"
     complement(6827..7816)
     /gene="kpsP"
     /codon_start=1
     /transl_table=11
     /product="GutQ/KpsF family sugar-P isomerase"
     /protein_id="BAA98732.1"
     /db_xref="GI:8978897"
     /translation="NPSPMISTDVCCQDILKQKEAVDFFQAFQPEAKMLAEKILGH
SGWFFPGVGGKGCVARLKAVATLQLSERALFFSPVLLHGLDGLVSPGDIYCLFSSKS
GETQLLDTVPHLKSRRAILVAITSPYNSNLAALSDLVILPSVAELDFNLIPNST
TCQMIFGDFLAMLFLHRSGLSTYKQHPSCQVGMKANGKVDMPFKTEYVFFCHLG
DKVSTSLVFSAYGCCVCIVDPQFLMGIFTDGLRSLASGYGGEVLSLSEKVMTA
NPRCTEDSDIAIALQIMESSPVAVLVPLDNEENRHVTGLLHMHTLAKAGLL"
     complement(7843..9024)
     /gene="sucB.2"
     complement(7843..9024)
     /gene="sucB.2"
     /codon_start=1
     /transl_table=11
     /product="dihydrolipoamide succinyltransferase"
     /protein_id="BAA98733.1"
     /db_xref="GI:8978898"
     /translation="MIFFRPPKIGTSGGSIIVMLKNLGHVARDPELIEVSTDKI
ATELPSKAGRLVRFVNGEVDGVLGILEEISEADDESTCPLTSCETKSEA
GSSSSWFSFAPVLAQREGIGLDNLQIAGTGGKRVTRDILRAYISEQOVSIEP
IQQEVNRIPLMSPLRAIASSLSKSDVPHASIVDVVDLMLNLSGERQFLDTH
GVKLTITFIVCLAQTRQFFPLNGSLDGTIIVMKSVGVAVNLAKEGVYVPIH
NQDQRLTYSIAKALADLSRAPLNKLDPSEVDQSVTVTNFGMTGALIGMPIRIYPIEV
ALIGTGTQKRVVVRDDDSLAKRMVYVTLTFDHRVLIDGIYGSSELTSLKKNLESVTM
G"
     complement(9038..10282)
     /gene="gltT"
     complement(9038..10282)
     /gene="gltT"
     /codon_start=1

```


RESULT 7	AR310754_08/c	Sequence split into 13 fragments	LOCUS AR310754	Accession AR310754
WPCOMMENT				
		Fragment Name	Begin	End
		AR310754_00	1	110000
		AR310754_01	100001	210000
		AR310754_02	200001	310000
		AR310754_03	300001	410000
		AR310754_04	400001	510000
		AR310754_05	500001	610000
		AR310754_06	600001	710000
		AR310754_07	700001	810000
		AR310754_08	800001	910000
		AR310754_09	900001	1010000
		AR310754_10	1000001	1110000
		AR310754_11	1100001	1210000
		AR310754_12	1200001	1230025
Continuation (9 of 13) of AR310754 from base 800001 (AR310754 Sequence 1 from patent US				
Query Match 99.1%; Score 1836.2; DB 6; Length 110000;				
Best Local Similarity 99.8%; Pred. No. 0;				
Matches 1849; Conservative 0; Mismatches 3; Indels 1; Gaps 1;				
QY	1	ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	60	
DB	28394	ATGAGTCTGGCAGATAAGCTGGTATTGCTTCTAGTAAACAGCTCGTCTTCTACTAGCAGA	28335	
QY	61	TCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCCCTCCCTCCACCCACGTTTGAT	120	
DB	28334	TCTGCAGACGTGGACTCAACGACAGCGACCGCACCTACGCCCTCCCTCCACCCACGTTTGAT	28275	
QY	121	GAATTAAAGACTCAAGCGCAACACAGCTTACGATATCTATCTTAACCTCAACATCACTAGCT	180	
DB	28274	GAATTAAAGACTCAAGCGCAACACAGCTTACGATATCTATCTTAACCTCAACATCACTAGCT	28215	
QY	181	GACATACAGCTGCTTTGTGTGAGCTCCAGGATGCTGCTCACTAATAATAAGGATACAGCG	240	
DB	28214	GACATACAGCTGCTTTGTGTGAGCTCCAGGATGCTGCTCACTAATAATAAGGATACAGCG	28155	
QY	241	GCTACTGATGAGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATCCCGATGCAGTT	300	
DB	28154	GCTACTGATGAGAAACCGCAATCGCTGGGAGTGGGAACTAAGAAATCCCGATGCAGTT	28095	
QY	301	AAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCAATCTT	360	
DB	28094	AAAGTTGGCGCGCAATTAACAGAAATAGCGAAATATGCTTCGGATTAACCAAGCAATCTT	28035	
QY	361	GACTCTTTAGGTAACTGACTTCTTTCGACCTCTTACAGCTGCTCTTCTCCAATCTGTA	420	
DB	28034	GACTCTTTAGGTAACTGACTTCTTTCGACCTCTTACAGCTGCTCTTCTCCAATCTGTA	27975	
QY	421	GCAACAATAACAAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA	480	
DB	27974	GCAACAATAACAAAGCAGCTGAGCTTCTTAAGAGATGCAAGATAACCCAGTAGTCCCA	27915	
QY	481	GGGAAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACATGCTTACAGGCACAG	540	
DB	27914	GGGAAACCGCTGCAATGCTCAATCTTTAGTTGATCAGACATGCTTACAGGCACAG	27855	
QY	541	ATAGAGAAAGTGAATCGATTAGGATGCAATATTTTGCAGGACAGAACCGTAGTGA	600	
DB	27854	ATAGAGAAAGTGAATCGATTAGGATGCAATATTTTGCAGGACAGAACCGTAGTGA	27795	
QY	601	GCTGTAGAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACAGTAAAGCAGCA	660	
DB	27794	GCTGTAGAAATGCTAAATCTAATAACAGTATTAAGCAACATAGATTACAGTAAAGCAGCA	27735	
QY	661	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAAT	720	
DB	27734	ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAGTTCCCGGACTCTCCAAAT	27675	
QY	721	CTTCAAGAACGGNACAAATGGTAAATACAGCTCAGAAAGATCTTAAATAATACAAACCT	780	

DB	27674	CTTCAAGAACGGNACAAATGGTAAATACAGCTCAGAAAGATCTTAAATAATACAAACCT	27615	
QY	781	GCAGATGGTCTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGCTCCAGCAACAGGA	840	
DB	27614	GCAGATGGTCTCTGATGTTCCAAATCCAGGAACCTACAGTTGGAGCTCCAGCAACAGGA	27555	
QY	841	AGTAGTATGCTAGTATTTCGTGTTTCCATGCTGTTAGTATGCTGAAAATGAGACCGCT	900	
DB	27554	AGTAGTATGCTAGTATTTCGTGTTTCCATGCTGTTAGTATGCTGAAAATGAGACCGCT	27495	
QY	901	TCCATTTTGTCTCTGGGTTTCGTGATGATTCACATGTTCAATACGAAAAATCTCTGAT	960	
DB	27494	TCCATTTTGTCTCTGGGTTTCGTGATGATTCACATGTTCAATACGAAAAATCTCTGAT	27435	
QY	961	TCTCAAGCTGCCAAACAGAGAGCTCGACGACAAAGCTAGACGAGAAAGCGCTGGAGAT	1020	
DB	27434	TCTCAAGCTGCCAAACAGAGAGCTCGACGACAAAGCTAGACGAGAAAGCGCTGGAGAT	27375	
QY	1021	GACAGTGTCTGCTGAGCGCTCGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTCTAGGTAAA	1080	
DB	27374	GACAGTGTCTGCTGAGCGCTCGCAGATGCTCAGAAAAGCTTTAGAAAGCGGCTCTAGGTAAA	27315	
QY	1081	GCTGGCAACCAACAGAGGCTACTCAATGCTTTTGAAGACAGATCGCTTCTGCTGCTG- TTGT	1139	
DB	27314	GCTGGCAACCAACAGAGGCTACTCAATGCTTTTGAAGACAGATCGCTTCTGCTGCTG- TTGT	27255	
QY	1140	GAGCGCAGAGTTCCTCCCGCTGAGCAAGTTCCTATAGGTCATCTGTAAGCAAGCTTTA	1199	
DB	27254	GAGCGCAGAGTTCCTCCCGCTGAGCAAGTTCCTATAGGTCATCTGTAAGCAAGCTTTA	27195	
QY	1200	CAAGACCTCAAAATCTACAGCTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGC	1259	
DB	27194	CAAGACCTCAAAATCTACAGCTTCTGATTATAAAACACAGATATCAGCAGGTTATGATGC	27135	
QY	1260	TTCAAAATCCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGATCTGATGATGAT	1319	
DB	27134	TTCAAAATCCATCAATGATGCTATGTTAGGCGCAAGAAATGATGCGATCTGATGATGAT	27075	
QY	1320	AAACATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGAGCACCAACAGAAAGCTCG	1379	
DB	27074	AAACATGTAAGTACCCCGCTCTCACAGATCCGTTCTAGAGCACCAACAGAAAGCTCG	27015	
QY	1380	AGGACCAAGAAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCT	1439	
DB	27014	AGGACCAAGAAAACACAGATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCT	26955	
QY	1440	TGGAGATGCTATAGTCAAGTTTCGGCACTACAAATCTGTAATGCGAGATCCTCAGTCGAA	1499	
DB	26954	TGGAGATGCTATAGTCAAGTTTCGGCACTACAAATCTGTAATGCGAGATCCTCAGTCGAA	26895	
QY	1500	TCCTCAAGGGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCC	1559	
DB	26894	TCCTCAAGGGAATAATGAGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCCCTCC	26835	
QY	1560	ACAGTTTGGCTATCCTTATGTCAACTTTCTAATGACTCTACACAGAAAGTTTCATAGCTAA	1619	
DB	26834	ACAGTTTGGCTATCCTTATGTCAACTTTCTAATGACTCTACACAGAAAGTTTCATAGCTAA	26775	
QY	1620	ATTAGAAAGTTTGTGCTGAAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTCCTT	1679	
DB	26774	ATTAGAAAGTTTGTGCTGAAAGGATCTAGGACAGCAGCTGAAATAAAGCACTTTCCTT	26715	
QY	1680	TGAACGAGCTCCTTGTGTTTATTCAGCAGTGTGCTCAATATCGCTCTCTATATTCGG	1739	
DB	26714	TGAACGAGCTCCTTGTGTTTATTCAGCAGTGTGCTCAATATCGCTCTCTATATTCGG	26655	
QY	1740	TTATCTCCAAATAACCAACCTAAGTGTGTTGGAGAGATTATTATGCTGCTTTCGTAAG	1799	
DB	26654	TTATCTCCAAATAACCAACCTAAGTGTGTTGGAGAGATTATTATGCTGCTTTCGTAAG	26595	
QY	1800	GCCTTTGTTGGGCTTACCAACACACTAGAACGATTTCTCAATAATAAAGA 1852		

[illegible]


```
QY 1741 TATCTCCAAATAA 1752
Db 1945 TATCTCCAAATAA 1956

RESULT 9
AX349491
LOCUS AX349491 1956 bp linear PAT 06-FEB-2002
DEFINITION Sequence 14 from Patent WO0202606.
ACCESSION AX349491
VERSION AX349491.1 GI:18615352
KEYWORDS Chlamydomphila pneumoniae
SOURCE Chlamydomphila pneumoniae
ORGANISM Chlamydomphila pneumoniae
REFERENCE 1
AUTHORS Ratti G. and Grandi G.
TITLE Immunisation against Chlamydia pneumoniae
JOURNAL Patent: WO 0202606-A 14 10-JAN-2002;
Chiron S.p.A. (It)
FEATURES
    Location/Qualifiers
        1..1956
            /organism="Chlamydomphila pneumoniae"
            /mol_type="unassigned DNA"
            /db_xref="taxon:83558"

ORIGIN
Query Match 94.4%; Score 1748.8; DB 6; Length 1956;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1750; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAGCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
Db 205 ATGAGCTGGCAGATAAGCTGGGTATGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 264

QY 61 TCTGCAGAGCTGGACTCAACGACAGCGGACCGACCTACGCTCCCTCCACCCACGTTGAT 120
Db 265 TCTGCAGAGCTGGACTCAACGACAGCGGACCGACCTACGCTCCCTCCACCCACGTTGAT 324

QY 121 GATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACTAGCT 180
Db 325 GATTATAGACTCAAGCGCAACAGCTTACGATATCTTTACCTCAACATCACTAGCT 384

QY 181 GACATACAGCTGCTTTGTGTGACCTCCAGGATGCTGTCACTAAATATAAGGATACAGCG 240
Db 385 GACATACAGCTGCTTTGTGTGACCTCCAGGATGCTGTCACTAAATATAAGGATACAGCG 444

QY 241 GCTACTGTATGAGGAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCGATT 300
Db 445 GCTACTGTATGAGGAACCGCAATCGCTGCGGAGTGGGAACTAAGAAATGCCGATGCGATT 504

QY 301 AAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATAAACCAAGCGATTCTT 360
Db 505 AAAGTTGGCGCGCAAAATACAGAAATAGCGAAATATGCTTCGGATAAACCAAGCGATTCTT 564

QY 361 GACTCTTTAGGTAACTGACTCTCTTCGACCTCTTACGCTGCTCTTCTCCAACTCTGA 420
Db 565 GACTCTTTAGGTAACTGACTCTCTTCGACCTCTTACGCTGCTCTTCTCCAACTCTGA 624

QY 421 GCACCAATTAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db 625 GCACCAATTAACAAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 584

QY 481 GGGAAAACGCTCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
Db 685 GGGAAAACGCTCAATTTGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 744

QY 541 ATAGACAAAGATGGAATGGGATAGGATGCAATTTTGCAGGACAGAAACGCTAGTGA 600
Db 745 ATAGACAAAGATGGAATGGGATAGGATGCAATTTTGCAGGACAGAAACGCTAGTGA 804

QY 601 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTCAGCTAAAGCAGCA 660
Db 1741 TATCTCCAAATAA 1752
Db 1945 TATCTCCAAATAA 1956

805 GCTGTAGAAATGCTAAATCTAATAACAGTATAGCAACATAGATTGAGCTAAAGCAGCA 864
661 ATCGCTACTGTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGACTCTCCAAAT 720
865 ATCGCTACTGTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGACTCTCCAAAT 924
721 CTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAACCT 780
925 CTTCAAGAGCGGAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAACCT 984
781 GCAGATGTTCTGTATGTTCCAAATCCAGGAATACAGTTGAGGCTCCAAAGCAACAGGA 840
985 GCAGATGTTCTGTATGTTCCAAATCCAGGAATACAGTTGAGGCTCCAAAGCAACAGGA 1044
841 AGTAGTATTTGTAGTATTCGGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 900
1045 AGTAGTATTTGTAGTATTCGGTTTCCATGCTGTTAGATGATGCTGAAAAATGAGACCGCT 1104
901 TCCATTTTGTATGCTGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATACTCTGAT 960
1105 TCCATTTTGTATGCTGGGTTTCGTCAGATGATTCACATGTTCAATACGGAATACTCTGAT 1164
961 TCTCAAGCTGCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGGAAAGCGCTGGAGAT 1020
1165 TCTCAAGCTGCCAACAGGAGCTCGCAGCACAAAGCTAGAGCAGGAAAGCGCTGGAGAT 1224
1021 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1080
1225 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGCTCTAGGTAAA 1284
1081 GCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTCTGTTG 1140
1285 GCTGGGCAACAAACAGGGCATACTCAATGCTTTAGGACAGATCGCTTCTGCTCTGTTG 1344
1141 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCTATAGGGTCATCTGTAAAAACAGCTTTAC 1200
1345 AGCGCAGAGTTCCTCCCGCTGCAGCAAGTTCTATAGGGTCATCTGTAAAAACAGCTTTAC 1404
1201 AAGACTCAAAATCTACAGGTTCTGATTATAAAACAGATATACAGAGTTATGATGCT 1260
1405 AAGACTCAAAATCTACAGGTTCTGATTATAAAACAGATATACAGAGTTATGATGCT 1464
1261 TACAAATCCATCAATGATGCTATGTTAGGGCAGCAAAATGATCGGACTCTGTATGATA 1320
1465 TACAAATCCATCAATGATGCTATGTTAGGGCAGCAAAATGATCGGACTCTGTATGATA 1524
1321 AACAAATGAAGTACCCCGCTCTCACAGATCCGTTCTCTAGAGCAGCAAGAGCTCGA 1380
1525 AACAAATGAAGTACCCCGCTCTCACAGATCCGTTCTCTAGAGCAGCAAGAGCTCGA 1584
1381 GGACAGAAAAACAGATCAAGCCCTCGTAGGGTGAATTTCTGGCAATAGCAGACTCTT 1440
1585 GGACAGAAAAACAGATCAAGCCCTCGTAGGGTGAATTTCTGGCAATAGCAGACTCTT 1644
1441 GGAGATGCTATAGTCAAGTTTCGCACTTACAAATCTGTAATGCGAGTCACTCAGTCGAAT 1500
1645 GGAGATGCTATAGTCAAGTTTCGCACTTACAAATCTGTAATGCGAGTCACTCAGTCGAAT 1704
1501 CTTCAAGCGAAATTAATGAGGAGATCAGAAAAAGCTTACATCGGAGTGAAGAAGCTCGA 1560
1705 CTTCAAGCGAAATTAATGAGGAGATCAGAAAAAGCTTACATCGGAGTGAAGAAGCTCGA 1764
1561 CAGTTTGGTATCTTATGTCGAATTTCTAATGATCTTACACAGAGTTCTATAGCTAAA 1620
1765 CAGTTTGGTATCTTATGTCGAATTTCTAATGATCTTACACAGAGTTCTATAGCTAAA 1824
1621 TTAGAAAGTTTCTTGTGCTGAAGATCTAGGACAGCAGCTGAAATAAAAACAGCTTTCCCTT 1680
1825 TTAGAAAGTTTCTTGTGCTGAAGATCTAGGACAGCAGCTGAAATAAAAACAGCTTTCCCTT 1884
1681 GAAACGAACTCCTGTTTATTTCAGCAGGTGCTGCTCAATATCGGCTCTCTATTTCTGCT 1740
1885 GAAACGAACTCCTGTTTATTTCAGCAGGTGCTGCTCAATATCGGCTCTCTATTTCTGCT 1944
```



```

Db      1895 GAACCAACTCTGTTATTACGAGGTCGTGGTCAATATCGGCTCTATATTCTGTT 1944
QY      1741 TATCTCCCAATAA 1752
Db      1945 TATCTCCCAATAA 1956

RESULT 11
LOCUS   CHT76KDA                2148 bp    DNA        linear    BCT 17-MAR-1994
DEFINITION Chlamydia pneumoniae 76 kDa protein gene, complete cds.
ACCESSION L23821
VERSION   L23921.1 GI:435961
KEYWORDS  76 kDa protein.
SOURCE    Chlamydia pneumoniae
ORGANISM  Chlamydia pneumoniae
          Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 2148)
AUTHORS   Perez Melgosa,M., Kuo,C.C. and Campbell,L.A.
TITLE     Isolation and characterization of a gene encoding a Chlamydia
          pneumoniae 76-kilodalton protein containing a species-specific
          epitope
JOURNAL   Infect. Immun. 62 (3), 880-886 (1994)
MEDLINE   94156481
PUBMED    7509320
COMMENT   Original source text: Chlamydia pneumoniae DNA.
FEATURES
          Location/Qualifiers
            1..2148
              /organism="Chlamydia pneumoniae"
              /mol_type="genomic DNA"
              /db_xref="taxon:83558"
            1..2148
              /codon_start=1
              /transl_table=11
              /product="76 kDa protein"
              /protein_id="AA23117.1"
              /db_xref="GI:435962"
              /translation="MTKRYHAWVVEGILNRLPKQFVKSVVDWNTFVPSSTTEKA
ATNAMKYKYVQWQVLGHSQVYVEALSVPEDIGRGNLLKFWFAQIAAAYDP
WFIGTGLGRKAGDWSATVRYVEALSVPEDIGRGNLLKFWFAQIAAAYDP
KEANGFTNYKGFSAIYMGITDLSFRAYGKPKANDKLGSDFTRKFDLGIISAGF
KSNFKIKPNRLALIISSFRFIYFNKPKRLLPEFLLVNPIGPPIDETERTPPAD
LSAQLEASANKSBEARIGACAPKESKIDSVERMSILBSAVNALMSLADKLGLIA
SNSSSSTSRKADSDSTATATPPPTDPTKQATAYDITFTSLADQALVS
LQDVTNKDTEETAIAAEWETKNADAVKGAQITELAKYASDNQAILDSUGKL
TSFDLQALLQSVAANNKAELLKEMQDNVPVPGKTPAIAQSLVDQDATATQEXD
GNAIRDAYPAGQVAGVAKNSNNSISNIDSAKALATAKTQIAEAKKFPDPILO
EAEQWVIOAEKOLKNIKPADGSDVPNGTGVGSKQOQSSIGSIRVSMILLDAENETA
STLMSGFROMIHMENTENPDSCAAQOELAAQARAANKAGDDSDAALADAKALEAL
GKAGQQQGIILNALGQIASNAVVSAGVILPQQVL"
ORIGIN

Query Match      62.9%; Score 1165; DB 1; Length 2148;
Best Local Similarity 99.9%; Pred. No. 5.2e-266;
Matches 1176; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY      1 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTTAGTAAACAGCTGGTCTTCTACTAGCAGA 60
Db      973 ATGAGTCTGGCAGATAAGCTGGGTATGCTTCTTAGTAAACAGCTGGTCTTCTACTAGCAGA 1032
QY      61 TCTGCAGACGTGACTCAACGACGACGCGACCTACCGCTCCCTCCACCCACGTTTGAT 120
Db      1033 TCTGCAGACGTGACTCAACGACGACGCGACCTACCGCTCCCTCCACCCACGTTTGAT 1092
QY      121 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTTACCTCAACATCAGCTAGCT 180
Db      1093 GATTATAGACTCAAGCGCAACAGCTTACGATATCTATCTTTACCTCAACATCAGCTAGCT 1152
QY      181 GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCTCACTAATAATTAAGGATACAGCG 240
Db      1153 GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCTCACTAATAATTAAGGATACAGCG 1212

```

```

QY      241 GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATCGAGTT 300
Db      1213 GCTACTGATGAGGAAACCGCAATCGCTCGGAGTGGGAAACTAAGAAATGCCGATCGAGTT 1272
QY      301 AAAGTTGGCGCGCAAAATTACAGAAATTACGAAATATGCTTCGGATAAACCAAGCGGATCTTT 360
Db      1273 AAAGTTGGCGCGCAAAATTACAGAAATTACGAAATATGCTTCGGATAAACCAAGCGGATCTTT 1332
QY      361 GACTCTTTAGTAAACTGACTTCTTTCGACCTCTTTACAGGCTGCTCTTCTTCCAAATCTGTGA 420
Db      1333 GACTCTTTAGTAAACTGACTTCTTTCGACCTCTTTACAGGCTGCTCTTCTTCCAAATCTGTGA 1392
QY      421 GCACAAATAACAGACGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480
Db      1393 GCACAAATAACAGACGAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 1452
QY      481 GGGAAACGCTTGCATTAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
Db      1453 GGGAAACGCTTGCATTAATGCTCAATCTTTAGTTGATCAGACAGATGCTACAGCGACACAG 1512
QY      541 ATAGAGAAAGATGGAATGCGATTAGGATGATATTTTTCAGGACAGACAGCTAGTGA 600
Db      1513 ATAGAGAAAGATGGAATGCGATTAGGATGATATTTTTCAGGACAGACAGCTAGTGA 1572
QY      601 GCTGTAGAAATGCTTAATCTAATAACAGATTAAGCAACATAGATTTCAGCTAAACAGCA 660
Db      1573 GCTGTAGAAATGCTTAATCTAATAACAGATTAAGCAACATAGATTTCAGCTAAACAGCA 1632
QY      661 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGGACTCTCCAATT 720
Db      1633 ATCGCTACTGCTTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTTCCCGGACTCTCCAATT 1692
QY      721 CTTCAAGAAAGCGGAACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAAATATCAAACT 780
Db      1693 CTTCAAGAAAGCGGAACAAATGCTTAATACAGGCTGAGAAAGATCTTAAAAAATATCAAACT 1752
QY      781 GCAGATGTTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACAAGGA 840
Db      1753 GCAGATGTTCTGATGTTTCCAAATCCAGAACTACAGTTGGAGGCTCCCAAGCAACAAGGA 1812
QY      841 AGTAGTATTGCTAGTATTTCGTGTTTCCATGCTGTTTAGATGATGCTGAAAAATGAGACCGCT 900
Db      1813 AGTAGTATTGCTAGTATTTCGTGTTTCCATGCTGTTTAGATGATGCTGAAAAATGAGACCGCT 1872
QY      901 TCCATTTTATGCTCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATCTCTGAT 960
Db      1873 TCCATTTTATGCTCTGGGTTTCGTCAGATGATTCATGTTCAATACGGAATCTCTGAT 1932
QY      961 TCTCAAGCTGCCAACAGAGGCTCCGACACAAGCTAGACAGCGAAAGCCGCTGGAGAT 1020
Db      1933 TCTCAAGCTGCCAACAGAGGCTCCGACACAAGCTAGACAGCGAAAGCCGCTGGAGAT 1992
QY      1021 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA 1080
Db      1993 GACAGTGTCTGTCAGCGCTGGCAGATGCTCAGAAAGCTTTAGAAAGCGGCTTAGGTAAA 2052
QY      1081 GCTGGGCAACACAGCGCATCTCAATGCTTTAGACAGATGCTTCTGCTGCTGTTG 1140
Db      2053 GCTGGGCAACACAGCGCATCTCAATGCTTTAGACAGATGCTTCTGCTGCTGTTG 2112
QY      1141 AGCGCAGGAGTTTCTCCCGCTGCAGCAAGTTCTATAG 1177
Db      2113 AGCGCAGGAG-TCTTCCCGCTGCAGCAAGTTCTATAG 2148

RESULT 12
LOCUS   AX045137                2238 bp    DNA        linear    PAT 24-NOV-2000
DEFINITION Sequence 7 from Patent WO0066739.
ACCESSION AX045137
VERSION   AX045137.1 GI:11343736
KEYWORDS  Chlamydia pneumoniae
SOURCE

```

ORGANISM	Chlamydia pneumoniae	Db	1510	ATAGAGAAAGATGAAATCGGATTAGGGATGCAATATTTTCAGGACAGAACCGTAGTGA	1569
REFERENCE	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia	Qy	601	GCTGTAGAAAATGCTAAATCTATAAGCAAGTAAAGCAACATAGATTTCAGTAAAGCAGCA	660
AUTHORS	Murdin, A.D., Omen, R.P., Wang, J. and Dunn, P.	Db	1570	GCTGTAGAAAATGCTAAATCTATAAGCAAGTAAAGCAACATAGATTTCAGTAAAGCAGCA	1629
TITLE	i(chlamydia) antigens and corresponding dna fragments and uses	Qy	661	ATCGCTACTGCTAAGACACAAATAGCTGAGGCTCAGAAAAGTCCCGGACTCTCCAAAT	720
JOURNAL	Thereof	Db	1630	ATCGCTACTGCTAAGACACAAATAGCTGAGGCTCAGAAAAGTCCCGGACTCTCCAAAT	1689
Patent: WO 0066739-A 7 09-NOV-2000;	Aventis Pasteur Limited (CA)	Qy	721	CTTCAAGAGCGGAAACAAATGCTGAGGCTGAGAAAGATCTTAAATATCAAACT	780
Location/Qualifiers	1. .2238	Db	1690	CTTCAAGAGCGGAAACAAATGCTGAGGCTGAGAAAGATCTTAAATATCAAACT	1749
source	/organism="Chlamydia pneumoniae"	Qy	781	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTCAGGAGCTCCAGCAACAGGA	840
	/mol_type="unassigned DNA"	Db	1750	GCAGATGGTTCTGATGTTCCAAATCCAGGAACCTACAGTTCAGGAGCTCCAGCAACAGGA	1809
	/db_xref="taxon:83558"	Qy	841	AGTAGTATTTGGTAGTATTCGTTTCAGTCTGCTGATGATGATGCTGAAATGAGCCGCT	900
	/note="unnamed protein product"	Db	1810	AGTAGTATTTGGTAGTATTCGTTTCAGTCTGCTGATGATGATGCTGAAATGAGCCGCT	1869
	766. .2238	Qy	901	TCCATTTTGGTCTGCTGGTTTCCTCAGATGATTCATGTTCAATACGGAATTCCTGAT	960
	/transl_table=1	Db	1870	TCCATTTTGGTCTGCTGGTTTCCTCAGATGATTCATGTTCAATACGGAATTCCTGAT	1929
	/protein_id="CAC17330.1"	Qy	961	TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAGCTAGACGCGAAAGCGCTGGAGAT	1020
	/db_xref="GI:1134373"	Db	1930	TCTCAAGCTGCCCAACAGGAGCTCGCAGCAACAGCTAGACGCGAAAGCGCTGGAGAT	1989
	/translation="MVNPIGPGPIDEITERTPPADLSAQGLEASAAKSAEAOIRIAGAE	Qy	1021	GACAGTCTGCTGACGGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA	1080
	AKPKESKTSVERMSILRSNAVLADKLGIASSNSSSSTSRSDVDTATAPTP	Db	1990	GACAGTCTGCTGACGGCTGGCAGATGCTCAGAAAGCTTTAGAGCGGCTCTAGGTAAA	2049
	PPPTFDYKTAQATYDTFTSTSLADIQAALVSLQDAVTNIXDIAATDETAASW	Qy	1081	GCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG	1140
	ETKNADAVKGAQITELAKYASDNQAILDSLGLTSLFDLLQAALQSVANNKAEEL	Db	2050	GCTGGGCAACAAACAGGCACTACTCAATGCTTTAGGACAGATCGCTTCTGCTGTTGTG	2109
	KEMQNPVPGKTPAIQSGVDQTDATQIEKGNARIDRAYFAGQWASGAVNAKSN	Qy	1141	AGCGAGAGTTCTCTCCGCTCGCAGCAAGTTCCTATAG 1177	
	NSINIDSAKAAIATQIAAOKKFPSPILQEAQOMVIQAEOLKNIKPADGSDV	Db	2110	AGCGAGAG-TACTCCGCTCGCAGCAAGTTCCTATAG 2145	
	PNPTGTVGSKQSGSISIRVSLMLDAENETASILMSGFRQIMHFNENTPDQAA	Qy			
	QOELAAQARAAGADPSAAALADAKALEALGKAGQQQILNALGQIASAAVUSA	Db			
	GVLFQQLVLMIRARYQAYVEQKLI SEEDLNSAVDHHHHH"	Qy			
ORIGIN		Qy			
Query Match	62.7%; Score 1161.8; DB 6; Length 2238;	Db			
Best Local Similarity	99.7%; Pred. No. 3e-265;	Qy			
Matches 1174; Conservative 0; Mismatches 2; Indels 1; Gaps 1;		Db			
1	ATGAGTCTGGCAGATAAGCTGGTATTGCTCTAGTAACAGCTCGTCTTACTAGCAGA 60	Qy			
970	ATGAGTCTGGCAGATAAGCTGGTATTGCTCTAGTAACAGCTCGTCTTACTAGCAGA 1029	Db			
61	TCTGACAGCTGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCACGTTTGTAT 120	Qy			
1030	TCTGACAGCTGGACTCAACGACAGCGACCGACCTAGCGCTCTCCACCCACGTTTGTAT 1089	Db			
121	GATTATAGACTAAGCGCAACAGCTTACGATATCTTTTACTCTCAACATCACTAGCT 180	Qy			
1090	GATTATAGACTAAGCGCAACAGCTTACGATATCTTTTACTCTCAACATCACTAGCT 1149	Db			
181	GACATACAGGCTGTTTGGTCAAGCTCCAGGATGCTGTCATTAATAAAGGATACAGCG 240	Qy			
1150	GACATACAGGCTGTTTGGTCAAGCTCCAGGATGCTGTCATTAATAAAGGATACAGCG 1209	Db			
241	GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAACTAAAGATGCCGATGCGATT 300	Qy			
1210	GCTACTGATGAGAAACCGCAATCGCTCGGAGTGGGAACTAAAGATGCCGATGCGATT 1269	Db			
301	AAAGTTGCGCGCAATTTACAGATTACGAAATATGCTTCGGATAACCAAGCATCTTT 360	Qy			
1270	AAAGTTGCGCGCAATTTACAGATTACGAAATATGCTTCGGATAACCAAGCATCTTT 1329	Db			
361	GACTCTTTAGTAACTGACTCTCTTTCGACCTCTTACAGGCTGCTCTCTCAATCTGTA 420	Qy			
1330	GACTCTTTAGTAACTGACTCTCTTTCGACCTCTTACAGGCTGCTCTCTCAATCTGTA 1389	Db			
421	GCAACAATAAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 480	Qy			
1390	GCAACAATAAACAAGCAGCTGAGCTTCTTAAAGAGATGCAAGATAACCCAGTAGTCCCA 1449	Db			
481	GGGAAAACGCTGCAATTTGCTCAATCTTTAGTATCAGACAGATGCTTACAGGACACAG 540	Qy			
1450	GGGAAAACGCTGCAATTTGCTCAATCTTTAGTATCAGACAGATGCTTACAGGACACAG 1509	Db			
541	ATAGAGAAAGATGGAATGCGATTAGGATGCATATTTTTCAGGACAGAAACGCTAGTGA 600	Qy			

RESULT 13
AX045135
LOCUS
DEFINITION Sequence 5 from Patent WO0066739.
ACCESSION AX045135
VERSION AX045135.1 GI:11343734
KEYWORDS
SOURCE Chlamydia pneumoniae
ORGANISM Chlamydia pneumoniae
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1
AUTHORS Murdin, A.D., Omen, R.P., Wang, J. and Dunn, P.
TITLE i(chlamydia) antigens and corresponding dna fragments and uses
JOURNAL Thereof
Patent: WO 0066739-A 5 09-NOV-2000;
Aventis Pasteur Limited (CA)
Location/Qualifiers
1. .1456
/organism="Chlamydia pneumoniae"
/mol_type="unassigned DNA"
/db_xref="taxon:83558"
101. .>1452
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC17379.1"
/db_xref="GI:11343735"
/db_xref="RENTREMBL:CAC17379"
/translation="MVNPIGPGPIDEITERTPPADLSAQGLEASAAKSAEAOIRIAGAE
AKPKESKTSVERMSILRSNAVLADKLGIASSNSSSSTSRSDVDTATAPTP

PPPTDDYKTAQATYDTFTSTSLADIQAALVLSQDQAVTNKIKOTAATDEBTATAAEW
ETKNADAVKGAQITELAKYASDNQAIILDSLGKLTFSFLLQALIQSVANNKKAELL
KEMONDVVPVGTIPAIQSLVDQTDATQTEKQGNARIDAYFACONASGAVNAKSN
NSINISDAKAAIATKQIAEAOKFPDPSPILOEAEQWVICAEDKLNKIPADGSDV
PNPPTVGSQKQSSISGIRVSMLLDIAENETASILMSGFRQMHMENTENPDSQAA
QQLAAQAPAAKAGGDSAAAALADAQALENAHGRAGQQQSILNAGQIASAAVVA
G"

ORIGIN

Query Match 62.1%; Score 1151; DB 6; Length 1456;
Best Local Similarity 100.0%; Pred. No. 1,1e-262;
Matches 1151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 60
DB 305 ATGAGTCTGGCAGATAAGCTGGGTATTGCTTCTAGTAACAGCTCGTCTTCTACTAGCAGA 364
QY 61 TCTGCAGAGCTGGACTCAACGACAGCGACCGCACCTACGCCTCCTCCACCCAGTTTGTAT 120
DB 365 TCTGCAGAGCTGGACTCAACGACAGCGACCGCACCTACGCCTCCTCCACCCAGTTTGTAT 424
QY 121 GATTATAAGACTCAACGCGAACAAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT 180
DB 425 GATTATAAGACTCAACGCGAACAAGCTTACGATACCTATCTTTACCTCAACATCACTAGCT 484
QY 181 GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACPAATATAAGGATACAGCG 240
DB 485 GACATACAGGCTGCTTTGGTGAGCCTCCAGGATGCTGTCACPAATATAAGGATACAGCG 544
QY 241 GCTACTGATGAGGAACCCGCAATCGCTGGGAGTGGGAACCTAAGAACTCCGATCGAGTT 300
DB 545 GCTACTGATGAGGAACCCGCAATCGCTGGGAGTGGGAACCTAAGAACTCCGATCGAGTT 604
QY 301 AAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGATTCT 360
DB 605 AAAGTTGGCGCGCAAAATTACAGAAATTAGCGAAATATGCTTCGGATTAACCAAGCGATTCT 664
QY 361 GACTCTTTAGGTAAATGACTCTTCCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 420
DB 665 GACTCTTTAGGTAAATGACTCTTCCTTCGACCTCTTACAGGCTGCTCTTCCAAATCTGTA 724
QY 421 GCAACAATAACAAACAGCTGAGCTCTTAAAGAGATCAAGATCAACCAAGCTAGTCCCA 480
DB 725 GCAACAATAACAAACAGCTGAGCTCTTAAAGAGATCAAGATCAACCAAGCTAGTCCCA 784
QY 481 GGGAAACCGCTGCAATGCTCAATCTTAGTTGATCAGACAGATGCTACAGCGACACAG 540
DB 785 GGGAAACCGCTGCAATGCTCAATCTTAGTTGATCAGACAGATGCTACAGCGACACAG 844
QY 541 ATAGAGAAAGATGGAATGGATTAGGATGCAATTTTGCAGGACAGACGCTAGTGA 600
DB 845 ATAGAGAAAGATGGAATGGATTAGGATGCAATTTTGCAGGACAGACGCTAGTGA 904
QY 601 GCTGTAGAAAATGCTAAATCTAATACAGTATAGCAACATAGATTCAAGTAAAGCAGCA 660
DB 905 GCTGTAGAAAATGCTAAATCTAATACAGTATAGCAACATAGATTCAAGTAAAGCAGCA 964
QY 661 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT 720
DB 965 ATCGCTACTGCTAAGACACAAATAGCTGAAGCTCAGAAAAAGTTCCCGAGCTCTCCAAAT 1024
QY 721 CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAACCT 780
DB 1025 CTTCAAGAGCGGAAACAAATGGTAAATACAGGCTGAGAAAGATCTTAAAAATATCAACCT 1084
QY 781 GCAGATGGTCTGATGTTCCAAATCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAGGA 840
DB 1085 GCAGATGGTCTGATGTTCCAAATCAGGAACCTACAGTTGGAGGCTCCCAAGCAACAGGA 1144
QY 841 AGTAGTATTGGTAGTATTCGTGTTTCCATGCTCTTATAGATGATCTGAAATATGACCGCT 900
DB 1145 AGTAGTATTGGTAGTATTCGTGTTTCCATGCTCTTATAGATGATCTGAAATATGACCGCT 1204

QY 901 TCCATTTTGATGTCTGGGTTTCGTCAGATGATTCATGTTCAATACCGAAAAATCCTGAT 960
DB 1205 TCCATTTTGATGTCTGGGTTTCGTCAGATGATTCATGTTCAATACCGAAAAATCCTGAT 1264
QY 961 TCTCAAGCTGCCCAACACAGGAGCTCGCAGCACAAAGCTAGCAGCAGGAAAGCGCTGGAGAT 1020
DB 1265 TCTCAAGCTGCCCAACACAGGAGCTCGCAGCACAAAGCTAGCAGCAGGAAAGCGCTGGAGAT 1324
QY 1021 GACAGTGTCTGTCAGCGCTCGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGGTAAA 1080
DB 1325 GACAGTGTCTGTCAGCGCTCGCAGATGCTCAGAAAGCTTTAGAACGCGCTCTAGGTAAA 1384
QY 1081 GCTGGGCAACACAGGAGCTACTCAATGCTTTAGACAGATGCTTCTGCTGCTTGTG 1140
DB 1385 GCTGGGCAACACAGGAGCTACTCAATGCTTTAGACAGATGCTTCTGCTGCTTGTG 1444
QY 1141 AGCGCAGGAGT 1151
DB 1445 AGCGCAGGAGT 1455

RESULT 14
AE016994
LOCUS Chlamydomophila caviae GPIC section 1 of 4 of the complete genome.
DEFINITION Chlamydomophila caviae GPIC
ACCESSION AE016994 AE015925
VERSION AE016994.1 GI:29834116
KEYWORDS
SOURCE Chlamydomophila caviae GPIC
ORGANISM Chlamydomophila caviae GPIC
REFERENCE 1 (bases 1 to 304769)
AUTHORS Read,T.D., Myers,G.S., Brunham,R.C., Nelson,W.C., Paulsen,I.T., Heidelberg,J., Holtzaple,E., Khouri,H., Federova,N.B., Carty,H.A., Umayam,L.A., Haft,D.H., Peterson,J., Beanan,M.J., White,O., Salzberg,S.L., Hsia,R.C., McClarty,G., Rank,N.G., Bavoil,P.M. and Fraser,C.M.

TITLE Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae
JOURNAL Nucleic Acids Res. 31 (8), 2134-2147 (2003)
MEDLINE 22569155
PUBMED 12682364
REFERENCE 2 (bases 1 to 304769)
AUTHORS Read,T., Myers,G., Brunham,R., Nelson,W., Paulsen,I., Heidelberg,J., Holtzaple,E., Khouri,H., Federova,N.B., Carty,H., Umayam,L.A., Haft,D., Peterson,J., Beanan,M., White,O., Salzberg,S., Hsia,R.-C., McClarty,G., Rank,R., Bavoil,P. and Fraser,C.

TITLE Direct Submission
JOURNAL Submitted (29-OCT-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

FEATURES
source
1..304769
/organism="Chlamydomophila caviae GPIC"
/mol_type="genomic DNA"
/isolate="GPIC"
/db_xref="taxon:227941"
57..1052
/gene="hemB"
/locus_tag="CCA00001"
57..1052
/gene="hemB"
/locus_tag="CCA00001"
/note="similar to GB:U16784, SP:P33395, and PID:290384; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="delta-aminolevulinic acid dehydratase"
/protein_id="AAP04754.1"
/db_xref="GI:29834117"
/translation="MSSLLLRPRNRKTAARLVAETALLPODFICPFVKEGKN
ICEETESLPGVVRSIDILLKEIBRLCSLGLKAVILFPVPSDLKDAYGSSNPKN
ICKSIYEVKAPFDLCVLSIDALDPYTHGHGIMDQGEVLNDESVRFIGNIATLAE


```

/transl_table=11
/product="exodeoxyribonuclease V, gamma subunit"
/protein_id="AAP04761.1"
/db_xref="GI:29834124"

Query Match
Best Local Similarity 53.8%; Score 361.2; DB 1; Length 304769;
Matches 903; Conservative 0; Mismatches 733; Indels 42; Gaps 6;

Qy 104 CTCACCCACGCTTTGATGATTAATAGCTCAAGCGCAACAGCTTACGATCTACTTTA 163
Db 21975 CTCCTCAAACTTTTATGATGAAGCTAAGACGCAAGCAGAGTGCAGAACTCGCTGCAGA 22034
Qy 164 CCTCAACATCACTAGCTACACATACAGGCTGCTTTGGTGAGCCTCCAGATGCTGTCACTA 223
Db 22035 GTGCGACTTATGATCAGTTCAAGACCGCTTTACAGCAGCTGCAAGATGCTGAAAC 22094
Qy 224 ATATAAGGATACAGCGGCTACTGATGAGGAAACCGGAATCGCTCGGAGTGGGAACTA 283
Db 22095 AGATGGAGCAATTAGCTACTACTGATGCAGAAAAGCTACAGTTGCTACATGGAAAACGG 22154
Qy 284 AGAATGCCGATGCAGTTAAAGTTGGCGGCAAAATTACAGAAATTAGCGAAATATGCTTCGG 343
Db 22155 CTCCTGAGCGCAGAAAGTACGCTGGATACACITTAACGATTTGGTGTCTATTCTTACAG 22214
Qy 344 ATAAACCAAGCGATTCTTGACTCTTTAGGTAACTGACTTCCTTCGACCTCTTACAGGCTG 403
Db 22215 AGAACCAAGCTTCTTGAGGCAATAAAGACGACCTCGTCTATGGATCAGATTATGGGAG 22274
Qy 404 CTCCTTCTCAATCTGTAGCAACAATAACAAAGCAGCTGAGCTTCTTTAAAGAGATGGAAG 463
Db 22275 CTGCGCGCAAGTAGAACCATAAACAACACTGCTGAGGAGTTAATTAACAGATTGGAAG 22334
Qy 464 ATAACCCAGTAGTCCCGGAGAAACCGCTGCAATTTGCTCAATTTAGTTGATCAGACAG 523
Db 22335 A-----AGCTGGGTTAGCTATCTCTGTGATAGTAGACCTTGAGAAAGCAATTA 22382
Qy 524 ATGTACAGCGACACAGATAGAGAAAGATGGAATTCGGATTAGGATGCATATTTTGCAG 583
Db 22383 CAACCTCAGGAACCTCAGGTTACTGAATTAGCAGATGCTATATCCGAAGCTTATGCTGCGG 22442
Qy 584 GACAGACGCTAGTGGAGCTGTAGAAAATGCTAAATCTAAATACAGTATAGCAACATAG 643
Db 22443 GGAATAACAGTACCCGCGCTGTGGGCGCAGCAGCAGCAATAACACCCCGCAATATAG 22502
Qy 644 ATTCAAGTAAAGCAGCAATCGCTACTGTGACACACAAATAGCTCAAGCTCAGAAAAAGT 703
Db 22503 AAGCTTCCAAACAAACTATTGCAATGCAAAAAAGTCATAGAACGCTCTTAAACTTG 22562
Qy 704 TCCCGACTCTCCAAATCTTCAAGAGCGGAAACAAATGGTAAATACAGCTGAGAAAGATC 763
Db 22563 CTCAGATTCTCCGATACTCAAAAGCTGCTTTGAAAGAACAAACACAGCAGCAAAAGATA 22622
Qy 764 TTAATAATCAAACTGCAGATGCTTCTGATGTTCCAAATCCAGAAC-----812
Db 22623 TCCTCAATGTGAACCTAGTGTGTGATGTGCTGCTATCGGTGCTCGAGCTCTCTG 22682
Qy 813 -TACAGTTGGAGGCTCCCAAGCAACAGGAAGTAGTATTGGTAGTATTCGTGTTCCATGC 871
Db 22683 GTAGTGTGGGACTTCTCAAAATCGCGGTGCTACTTAGGGAAGTTGCGTATCGATGT 22742
Qy 872 TGTTAGATGATCTGAAATAGACCGCTTCCATTTTGTATGCTCGGTTTCGTAGATGA 931
Db 22743 TATTGACTGATGTTGATATGAATGAACCGCAGCGATCATTTATGCAAGTTTCAGAAATATGA 22802
Qy 932 TTCATATGTTCAATACGGAATTCCTGATTTCAAGCTGCCCAACAGGAGCTCGCAGCAC 991
Db 22803 TCGATAAATTCCTATGATCAAAATCTCTGATTTTACAGCGCTTTAGAGAGATTTATGAAT 22862
Qy 992 AAGCTTA-----GAGCAGCGGAAGCGCGTGGAGATCACAGTGTGCTGCAGCGCTGG 1042
Db 22863 AAGTAAACGACTTATCAACGCGAGATCAATCTCTGCAGATCGGAAGCTACAGCAACTAC 22922

```

```

Qy 1043 CAGATGCTCAGAAAGCTTTAGAACGGCTCTAGGTAAGCTGGSCAACAAACAGGGCTATC 1102
Db 22923 AAGAAATACAAACAAACCATCAAGATGCCCTTCAAGGGAAGCTCGCGTCAAGACGGCATGA 22982
Qy 1103 TCAATGCTTTAGCAGCAGATCGCTTCTGCTCTGTTGTGAGCGCAGGATTCCTCCCGCTG 1162
Db 22983 TCAATGCTTTAGGAGCTTATAACAAACAGCAGCTTCAATTTCTACAGGAGCTCCTATCGCTT 23042
Qy 1163 CAGCAAGCTTCTATAGGCTCATCTGTAAACAGCTTTTACAAGAC---CTCAAAATCTACAG 1219
Db 23043 CTGCAATCAAGGTGGATCAGCTGTAAGCAGCTTTTACAAACAGGATCTACTCTCGGA 23102
Qy 1220 GTTCTGATATPAAACACA---GATATCAGCAGGTTATGATGCTTACAAATCCATCATG 1276
Db 23103 GTTCTAAATCTTACGCGGATTCCTTATCTGCAAGGTATGGGCGATCAATCTTTTAAATG 23162
Qy 1277 ATGCTATGTTAGGCGCAGAAATGATCGACTCTGTATGTGTATAAACAATGTAAGTACCC 1336
Db 23163 ATGTGTTACTCAG---TAGTAGTGCACTTAACCGTGAGGTTTATAGTCGTACATCGATC 23219
Qy 1337 CCGCTCTCACACAGTCCGTTTCTTAGACAGCAACAGAAAGCTCGAGGACCAAGAAAAACAG 1396
Db 23220 CAGCATTAACGACAGCAGTTTCTAGAACAGAAACTCGGCTCGTGATAATGATAACGCAG 23279
Qy 1397 ATCAAGCCCTCGCTAGGCTGATTTCTGGCAATAGCAGAACTCTTGGAGATGCTTATAGTC 1456
Db 23280 CTCAGCGCTTTTGAAGAACTATAGCTCTAATAGTAATCTCTTGGGATGTTTATGCAAT 23339
Qy 1457 AAGTTTCGGCACTCAATCTGTATGAGATCACTCAGTGGAACTCCTCAAGCGAATAATG 1516
Db 23340 CCGTAGGTGTTATGCAAACTGCTAGTGTATTACAAATATCCCAACAGGATGAG 23399
Qy 1517 AGGAGATCAGACAAAGCTTACATCGGAGTGACAAAGCTCCACAGTTTGGCTATCCTT 1576
Db 23400 AAGAAATCAACAGACAGCTCACTCTGAGGTTACGAAAGCTCCGCACTCAGGTTATCCTC 23459
Qy 1577 ATGTGCAACTTTCTAATGACTCTACAGAGTTTCATAGTAAATTAGAAGTTTGTGTTG 1636
Db 23460 ATGTACAGCTTTCTAAGCACTCTACGAAAGATTCATGTCTCAACTCGAATGAATTTG 23519
Qy 1637 CTGAGGATCTPAGNACAGCAGCTGAAATAAAGCACTTTCCTTTGAAACGAACTCCTTGT 1696
Db 23520 TTCAGGATCGAAAGACTTCCGCAAGCAAAAGAGCTGCGTTTGAAACAGCCTTTGT 23579
Qy 1697 TTATTACAGAGCTCGTGGTCAATACGCTCTCTATATTCGTGTTATCTCCAAATACA 1754
Db 23580 TCATCCAGCAGTATTAGTGAACGATAGCATCTCTGTTCTCGGATACCTACAGTAAAA 23637

```

```

RESULT 15
BD264444
LOCUS
DEFINITION
    Compounds and methods for treatment and diagnosis of chlamydia
    infection.
ACCESSION
    BD264444
VERSION
    BD264444.1 GI:33074212
KEYWORDS
    JP 2002531129-A/32.
SOURCE
    Chlamydia trachomatis
ORGANISM
    Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE
    1 (bases 1 to 1537)
AUTHORS
    Probst,P., Bhatia,A., Skeiky,Y.A.W., Fling,S.P., Jen,S. and
    Stromberg,E.J.
TITLE
    Compounds and methods for treatment and diagnosis of chlamydia
    infection
JOURNAL
    Patent: JP 2002531129-A 32 24-SEP-2002;
CORIXA CORP
OS
    Chlamydia trachomatis
PN
    JP 2002531129-A/32
PD
    24-SEP-2002
PF
    08-DEC-1999 JP 2000586916
PR
    08-DEC-1998 US 09/208277,08-APR-1999 US 09/288594 PR
    01-OCT-1999 US 09/410568,22-OCT-1999 US 09/426571 PI PETER

```

```

PROBST, AJAY BHATIA, YASIR A W SKEIKY, STEVEN P FLING, PI SHYIAN
PI, ERICA JEAN STROMBERG
PC C12N15/09, C12N15/09, A61K35/12, A61K38/00, A61K39/118, A61K39/395,
PC A61K39/395,
PC A61K45/00, A61K48/00, A61P33/00, C07K14/295, C07K16/12, C07K19/00,
PC C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/68, G01N33/53, G01N33/PC
532,
PC
G01N33/569//C12N1/19, C12R1/01, C12N15/09, C12R1/01, C12Q1/68, PC
C12R1/01,
PC C12N15/00, C12N15/00, A61K37/02, C12N15/00, C12R1/01 CC
Compounds and methods for treatment and diagnosis of CC
chlamydial infection
FH key Location/Qualifiers
FT source 1..1537
FT 1..1537 /organism='Chlamydia trachomatis'.
FEATURES
source
1..1537
Location/Qualifiers
/organism='Chlamydia trachomatis'
/mol_type='genomic DNA'
/db_xref='taxon:813'
ORIGIN
Query Match 10.1%; Score 187.2; DB 6; Length 1537;
Best Local Similarity 53.0%; Pred. No. 1.6e-33;
Matches 480; Conservative 0; Mismatches 408; Indels 18; Gaps 3;
QY 863 TTTCATGCTGTAGATGCTGAAATGAGACCGCTTCATTTTGATGCTGGGTTTC 922
DB 186 TTTCCTTGTGCTTGTAGATGCTGAGCAATGAATGGCAGCGATTGCAATGCAAGGTTTC 245
QY 923 GTCAGATGATTCATGCTTCAATACGGAATCTGATTTCTCAAGCTGCCACAGAGC 982
DB 246 GATCTATGATCGAACAAATTAATGTAACAAATCTTGCAACAGCTAAAGAGCTCAAGCTA 305
QY 983 TCGCAGCACAAAGCTAGACGACGAAAGCCGCTGGAGATGACAGTGTGCTGCGCGCTGG 1042
DB 306 TGGAGGCTCAGCTCA---CTGCGATGTCAGATCAACTGGTGTGGGATGGCGAGCTCC 362
QY 1043 CAGATGCTCAAGAGCTTTAGACGGCTCTAGTAAAGCTGGGCAACAAC----- 1093
DB 363 CAGCCGAATACAGCAATCAAGATGCTCTTGGCAAGCTTTGAACAACCATCAGCAG 422
QY 1094 AGGCGATCTCAATGCTTTAGGACAGATCGCTTCTGCTGCTGTGTGCGCGCAGGATTC 1153
DB 423 ATGGTTTAGCTACAGCTATGGACAGTGGCTTTTGCGAGCTGCCAAGTTGGAGGAGCT 482
QY 1154 CTCGCGCTGCAGCAAGTTCTATAGGGTCACTCTGTAAACAGCTTTTACAAGACCTCAAAAT 1213
DB 483 CCGCAGGAACAGCTGGCAGCTGCCAGATGAATGTAAACAGCTTTACAAGACGCGTTT 542
QY 1214 CTACAGGTTCT-----GATTATAAACACAGATATCAGCAGGTTATGATGCTTACAAAT 1267
DB 543 CTTGCGACTTCTCCAGCTCTTTATGACGACGACCTTTCCGATGGATATTTCTGCTTACAAA 602
QY 1268 CCATCAATGATGCTATGTTAGGCGCAGGAATGATGGAGCTCGTGATGTATAAACAATG 1327
DB 603 CACTGAATCTTTATATTCGAAAGACAGAGCGCGTGCAGTCAGCTATTAGTCAAACTG 662
QY 1328 TAAGTACCCCGGCTCTCACAGATCCGTTCTTAGAGCACGAACAGAGCTCGAGGACCAG 1387
DB 663 CAAATCCCGGCTTTCCAGAGCGTTTCTCGTTCTGGCATAGAAAGTCAAGGACGCGAGTG 722
QY 1388 AAAAAACAGATCAAGCCCTCGTAGGGTGATTTCTGCAATAGCAGAACTCTTGGAGATG 1447
DB 723 CAGATGCTAGCAAAGACGACGAGAAACTATTGTGACAGATAGCCAAACGCTTAGGTGATG 782
QY 1448 TCTATAGTCAAGTTTCGGCACTCAATCTGTATGACAGATCACTCAGTCGAATCCTCAAG 1507
DB 783 TATATAGCGGCTTACAGGTTCTTGGATCTTTTGATGTCTACGATTTGTGAGCAATCCGCAAG 842

```

```

QY 1508 CGAATAATGAGGAGATCAGACAAAAGCTTACATCGGCAGTGACAAAAGCTCCACAGTTTG 1567
DB 843 CAATTCAGAGAAGATATGAGAAGCTCACGGCATCTATTAGCAAAAGCTCCACATTTG 902
QY 1568 GTATCCCTTATGTGCAACTTTCTAATGACTCTACACAGAAGTTCTAGTAAATTAGAAA 1627
DB 903 GGTATCCCTGCTGTTTCTGAGAAATCTGTGGATAGCTTGCAAGAGTTTGTGTCACAAATTGAAA 962
QY 1628 GTTTGTTTCTGAAGGATCTAGACAGCAGCTGAAATAAAAGCACTTTCTCTTTGAAAACGA 1687
DB 963 GAGAGTTTCTTATGTTGGGAACGTAGTCTGCGAGAATCTCAAGAGAATGGGTTTAGAAAAC 1022
QY 1688 ACTCCTTGTATTATCAGCAGGTGCTGTCATATATCGGCTCTCTATATTTCTGTTATCTCC 1747
DB 1023 AGCCCGCTTTCATTCAACAGAGTTGTTGTAACATTGCTTCTATTTCTCTGTTATCTTT 1082
QY 1748 AATAAC 1753
DB 1083 CTTAAC 1088

```

Search completed: March 24, 2004, 01:57:11
Job time : 7202.51 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	2844	99.8	651	2	D72042		conserved hypothet
2	2844	99.8	651	2	E86581		CHLPN 76 kDa homol
3	1870	65.6	715	2	I40729		hypothetical 76k p
4	843	29.6	647	2	G71490		hypothetical prote
5	210	7.4	2055	2	T31110		extracellular matr
6	206	7.2	6713	2	B89921		hypothetical prote
7	205	7.2	1832	2	S33441		EF protein - Strept
8	199	7.0	971	2	B90835		probable tail fibre
9	199	7.0	973	2	C85693		probable membrane
10	195.5	6.9	2271	2	F90073		hypothetical prote
11	189	6.6	1063	2	D86731		hypothetical prote
12	189	6.6	1365	2	T30822		LmpI protein - Myc
13	187.5	6.6	2481	2	D90011		FmtB protein [impo
14	187.5	6.6	4776	2	E95206		cell wall surface
15	179.5	6.3	641	2	C83206		methyl-accepting c
16	178.5	6.3	1156	2	T34852		probable secreted
17	176.5	6.2	232	2	T34434		hypothetical prote
18	175	6.1	1072	2	A86827		hypothetical prote
19	174.5	6.1	1104	2	S59310		probable membrane
20	174	6.1	1829	2	T24583		hypothetical prote
21	174	6.1	2541	2	S11661		talin - mouse
22	172	6.0	1306	2	S25370		MSB2 protein - yea
23	171	6.0	1122	2	G64887		probable tail fibre
24	170.5	6.0	1566	2	A43607		cell surface antig
25	168	5.9	628	2	F84219		HtrII transducer [
26	168	5.9	993	2	C90072		hypothetical prote
27	168	5.9	2155	2	AD2742		conserved hypothet
28	168	5.9	2155	2	C97523		hypothetical prote
29	167.5	5.9	891	2	G41662		130kDa surface exclu

C;Accession: G71490
R;Stephens, R.S.; Kalman, S.; Lammell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A;Reference number: A71570; MUID:9900809; PMID:9784136
A;Accession: G71490
A;Molecule type: DNA
A;Residues: 1-647 <ARN>
A;Cross-references: GB:AE001333; GB:AE001273; NID:g3329068; PIDN:AAC6826.1; PID:g332906
A;Experimental source: serotype D, strain UW-3/Cx
C;Comment: This sequence was originally identified as homologous to part of a sequence (PIR:E72042).
C;Genetics:
A;Gene: CT622

Query Match 29.6%; Score 843; DB 2; Length 647;
Best Local Similarity 35.5%; Pred. No. 2.3e-35;
Matches 213; Conservative 113; Mismatches 243; Indels 31; Gaps 14;

QY 1 MSLADKLGIASSSSSTSSRS-ADVDSTTATPTPPPTFFDDYKTAQAYDTTFTSTSL 59
DB 60 TSLRD--AAILKNSSPTDLSQLEASTSTVTRVAARDYNEAKSNFTAKSGLNATL 117
QY 60 ADIOAALVSLQDAVNIKD--TAIAAEWETKNADAVKVGCAQITELAKYASDNOA 118
DB 118 AEYETKMADLAALQDMERLAKQAEVTRKEALQEKQ----EVIDKLNQLVKLEKQNT 173
QY 119 ILDSLGKUTSFDLLQALLQSVANNKAAELKEMQDNPVPGKTPAQAQSLVDQTDATA 178
DB 174 LKETLTITDSDAQIPAINSOLEINKNSAQDIKLEGQNI---SYEAVLTNAGEVIKASS 230
QY 179 TQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAIATAKTQIAEAQK---F 235
DB 231 EAGIKLQALQSLVDAGQSOQAALVQAQNNSPDNIAATKLLIDAEETKVELKQEHGTL 290
QY 236 PDSPILOBAEOMVIAEQKDLNKPAGDSVDPNPGTTVGGSKQGGSSITGSI-----RV 288
DB 291 TDSPLVKAEBQISQAQDKIQEIKES-GSDPIVGFPS-GSAASAGSAGALKSSNNSGRI 348
QY 289 SMLDDAENETASILMSGFROMIHMFTENPDSQAQQLAAQAPAAK---AAGDSDAAA 345
DB 349 SILLDDVDNEMAAIAMQFRSMIEQFNVNNPATAKELQAMEAQLTAMSDQLVGADGELPA 408
QY 346 ALADAQKALEALGKAGQOQGLNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKT--S 403
DB 409 EIQAIKDLAAL-KQPTDGLATAMGQVAAAKVGGSGAGTAGTQVQNVKQLYKTAFS 467
QY 404 KSTGSDYKTSAGYDAYKSIDAYGRANDATRDVINNVSTPALTTRSVPRARTEARG-P 462
DB 468 STSSSSYAAALSDGYSAYKTLNLSYSESRG-VQSAISQTANPALSRVSRSIGESQGRS 526
QY 463 EKTQALARVTSNGRTLDGVYSQVSALQSVWQITQSNPQANNEIROKLTSAVTKPQPF 522
DB 527 ADASQAAETIVRDSQTLGDVYSRLQVLDLSLSTIVSNPQVQNEEIMOKLTASIKAPQF 586
QY 523 GYPVQLGNDSTQKFIKLESIFARSGRTAAEIKALSFETNSLFTQQVLNIGLSYGYL 582
DB 587 GYPVQNSADSLQFAAQGLREFVDGERSLAESENAPRKQPAFIQQVLNIGLSYGYL 646

RESULT 5
T31110
extracellular matrix binding protein - Abiotrophia defectiva (fragment)
C;Species: Abiotrophia defectiva
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C;Accession: T31110
R;Manganelli, R.; van de Rijn, I.
Infect. Immun. 67, 50-56, 1999
A;Title: Cloning and characterization of emb, a gene encoding the major adhesin of Streptococcus pneumoniae
A;Reference number: Z20988; MUID:99081722; PMID:9864195
A;Accession: T31110
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-2055 <MAN>
A;Cross-references: EMBL:AF067776; NID:g3249002; PID:g3249003; PIDN:AAD03320.1
C;Genetics:
A;Gene: emb

Query Match 7.4%; Score 210; DB 2; Length 2055;
Best Local Similarity 21.8%; Pred. No. 0.0098;
Matches 129; Conservative 90; Mismatches 262; Indels 112; Gaps 24;

QY 10 ASSSSSSSTSSRSADVDSTTATPTPPPTFFDDYK----- 43
DB 1381 ATSNDDETAQKNEGTQAINAVPQIPKAKT--DAKNVTAQADRKDAIENDPNLTREKV 1438
QY 44 -----TQATAYDTTFTSTSLADIAQALVSLQDAVNIKD--TAATDETAIAAEWET 94
DB 1439 AAKAKVDAEAKKAKDAIDAATSNADVTAKQNEGTKAINDVPTPTAKTDKAKVQRAIDA 1498
QY 95 KNADAVKVGCAQITELAKYASDNOAILDLSGLTSPDILLQALLOS--VANNKAAELIKE 152
DB 1499 KK-DAIEKDPNLTREKDA--KAKVDAEAKKAK-DAIDAATSNADVTAKQNEGTKAIND 1554
QY 153 MQDNPVPGKTPAQAQSLVDQTDATATQIEKDGNAIRDAYFA-----QGNASGAVEN 204
DB 1555 VPQTPET--AKTDA-KNAVTAQADAKKDAIEKDPNLTREKDAKAKVDAEAKKAKDAIDA 1611
QY 205 AKSN-----NSISNIDSAKAIATAKTQIAEBAQK-KFPDPSILOBAEOMVI 249
DB 1612 ATSNADVTAKQKGNAINAVPQTPPTAKTDKAKNAVTAQADAKKDAIENDANLTREKDA 1671
QY 250 QAEKDLNKPAGDSVDPNPGTTVGGSKQGS-SIGST-RVSMELDDAEN--ETASILM 304
DB 1672 KAKVDAEATKAKNALDAATSNADVTAKQNEGTKAINDVPTPTAKTDKAKVDAQATDKK 1731
QY 305 SGFRQIMHMTEN-----PDSQAQQLAAQAAKAAAGDSDAAAALADAQKALEAAL 358
DB 1732 SAI-----ENDPALTREKDAKAKVDAEATKAKNAIDAATSNADVTAKQ- 1776
QY 359 GKASQOQGLNALGQIASAAVVSAGVPPAAASSIGSSVKQLYKTSKSTGSDYKTIQISAGY 418
DB 1777 -DAG--KNAINAVPQTPPTAKTDKAKNAVDAQATDKKAAIENDPALTREKDAKAKVDAEA 1833
QY 419 DAYKSIDAYG-----RARNDATERDVINNV-STPA-----LTRSVPRARTEARGPEKTDQA 468
DB 1834 KKAKDAIDAATSNADVTAKQKGDALNAVPTPTAKTDKAKNAVDAQATDKKSAIENDPA 1893
QY 469 LARVISGNSRTLDGVYSQVSALQSVWQITQSNPQANNEIROKLTSAVTKPQPF 521
DB 1894 LTR--EEKDAVKAKVDAEAKKAKDAIDAATSNADVTAKQTEG--TQAINAVPQ 1942

RESULT 6
B89921
hypothetical protein ebhA [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: B89921
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A;Reference number: B89758; MUID:21311952; PMID:11418146
A;Accession: B89921
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-6713 <KUR>
A;Cross-references: GB:BA000018; PID:g13701232; PIDN:BAB42527.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: ebhA

Query Match 7.2%; Score 206; DB 2; Length 6713;
Best Local Similarity 21.5%; Pred. No. 0.074;

Matches 145; Conservative 97; Mismatches 279; Indels 154; Gaps 31;

QY 5 DKLGIASSNSSSSRSADVDS---TTATATPTPPPTFDYDKTAQAYDTIFTS---TSL 59
Db 2215 EKLGAAAKQTAKSDIGRLTDLNNAORTAANAEDQAPNL-----AAVTAANKKATSLMTAM 2269
QY 60 ADIQAAVLVLQDAVNIK-----DTAATDEBETAIAAEWETKNADAVKVGQAQIT 107
Db 2270 GNLKHALAEKDNTRSVTADQPKQQAQYDVTVAEAITNA--NGSNANETQVQAALN 2327
QY 108 ELAKYASDNQAIDLSGLKTSFDLLQALLOSVAANNKKAELLKEMO-DHPVPEKTPAI 166
Db 2328 QLNQAKND-----LNGDNKVAQAKETAKRALASYGNLNAQSTAAATSDIDNATTVADVTA- 2382
QY 167 AQLVDDTDTATATQIE---KDGNAI-----RDAYF--AGNAGAVENAKSNN 209
Db 2383 AQNTANELNTAMGQLQNGINDQNTVKKQVNFDTADQKQKDAYTNVAVNAQOGLDKANGON 2442
QY 210 SISNIDSAKAAIATAKTQIAEAAQKFPDPSILOBAEQWVIAEKD-----LKNIK 259
Db 2443 -----MTKAQVEAALNQVTTAKNAL-----NGDANVRQAQSDAKANLGLTLHLNNAQ 2489
QY 260 PADGSDVNPFTTGG-----SKQGG-----SSIGS---IRVSMILLDAENETASI 302
Db 2490 KQDUTSQIEGATTGVNSVKTAKQDLGAMQRLSAIAANKDQTKASENYIDADPTKKTA 2549
QY 303 LMSGFRQMIHMFNTE---NPDQAQOQLAAQARAQAAKAA--GDDSAALADADAQKALEAAL 358
Db 2550 FDNALITQAESVYNKDHGNTKDKQAVEQAIQSVSTENALNGDANLQCAKTEATQAINLT 2609
QY 359 GKAGQQGILNALQCIASAAVVSAGVPPA--AASIGSSVKQLYKTSKSGDYKTOISA 416
Db 2610 QLNTFPK---TALKQOVNAAQVRSVGTDLKNSATSLNAMDQL---KQATGDHDTTVAG 2662
QY 417 GY-----DAYKSINDAGARN--DATRVINNVV--TPALTRSVPRARTBEARG----- 461
Db 2663 GNYTNASPDKQCAVTDAYNAAGNIVGSPNVTNAQVTAATQORNNAAETSLNGDNTLAT 2722
QY 462 -PEKTDQALRV-----ISGNSRTLGDVYSQVLSALQVMQITQS-----NP 501
Db 2723 AKQQAQKALQWTHLSDAQKQSIQGI-----DSATQVTGQSVKDNATLNDNAMNQLRNS 2778
QY 502 QANNEBIRQKLTSAVTRPQPGYPPVQLSNDSTOKF---IAKLESLEFAEGSRTAABEIKAL 558
Db 2779 IANKDEVKA-----SQPYVDADTDKQAYNTAVTSAENIINATSQPTLDPSAV 2826
QY 559 SPFETNSLFIQOVLVN 573
Db 2827 TQAAQNVNTKALN 2841

RESULT 7
EF protein - Streptococcus suis
C:Species: Streptococcus suis
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: S33441
R:Smith, H.B.; Reek, F.H.; Vecht, U.; Gielkens, A.L.J.; Smits, M.A.
A:Description: Repeats in an extracellular protein of weak-pathogenic strains are absent
A:Reference number: S33441
A:Accession: S33441
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1822 <SMI>
A:Cross-references: EMBL:X71880; NID:g298031; PIDN:CAA50714.1; PID:g298032

Query Match 7.2%; Score 205; DB 2; Length 1822;
Best Local Similarity 23.7%; Pred. No. 0.015;
Matches 145; Conservative 82; Mismatches 229; Indels 156; Gaps 27;

QY 37 PTFDYKTKTAQ-----TAYDTITSTSLADIQAALVSLQDA--VTNIKDTAA 81

Db 1265 PNLSDQKQAEIKKLTDAVAKTLATIRDNADKETQEAQKAAQALADLEKAKETOKIADKAA 1324
QY 82 TDEETAIA--AEWETKNADAVKVGQAQITELAKYASDNQAIDLSGLKTSFDLLQAAALQOS 139
Db 1325 IDRLTTLIVKDGELKATQKQAKIAKAAAAKAAIAISNPNLTDAEKKTTFTDAVDA---E 1380
QY 140 VANNKA-----AELLKEMQNDPVVPGKTPAIAQSLVDQTDATA--TOIEKDGNAIRDA 191
Db 1381 VAKANDAIQAATSPADVQKE--EDAGVA-----AIAEDVLDAAKQDAKAKIAKAAAAKEA 1434
QY 192 YFAGQWASGA-----VENAKSNSISNIDSAAKAIATAKTQIAEAAQKFPDPSIL 241
Db 1435 IGSNPNLTDAEKKTTFTDAVDAEVAKANDAI-----AATSPADVQKE--EDAGVA 1482
QY 242 QEAQOVVIOAEKDLKNIKADGSDVNPVPGTGVGSGKQOGSSIGSIRVSMILLDAENETAS 301
Db 1483 AIAEDVLDAAKQDAK--KIAKESD-----AAKSAIDANPNLTDAEKESAK 1526
QY 302 IMSGFRQMIHMFNTPNPSQAAQQLAA-----QARAQAAAGDDSAALADAQK----- 352
Db 1527 -----KAVDADAKAATDAIDASTSPVEAQAEDKGVGSIQAQDVLDAKQDAK 1573
QY 353 ---ALEANLGA-----CQOQGILNALQCIASAA--VVSAGVPPAAASS-----I 392
Db 1574 NKTAKEVAAAKEAIDANPNLSDAEKESAKKAVDADAKATTDAIDASTSPVEAQAEDKGV 1633
QY 393 GSSVKQLYKTSKSGDYKTOISAGYDVAKYSDINAYGR---ARNDATRDVIN---NVSTP 446
Db 1634 GSIRQDVLDAK--QDAKNKIAKESDAKSAIDANPNLTDAEKESAKKAVDADAKAATD 1690
QY 447 ALTRSVPRARTBEARGPE-KTDQALARVI-----SGNSRTLGDVYSQVLSALQVMQITQS 499
Db 1691 AIDAST--SPVEAQAEDKGVGAIAKDIILDAKQDAKAKIAKESAKSVIDSNPNLTDA 1748
QY 500 NPQANNEIRQKLTSAVT-----KPPQFGVPPYVQLSNDSTOKF 537
Db 1749 AKAAKSEIDKAVEAIVLNGVTRTQVELEKILPAAALIKPAKVTVPVD--PNNLTEKE 1807
QY 538 IAKLESLEFAEGS 549
Db 1808 IARIKAFLEKENN 1819

RESULT 8
B90835
probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain R1:
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: B90835
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B90835
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA035073.1; PID:gl3361114; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 050952
A:Genetics:
A:Gene: ECs1650

Query Match 7.0%; Score 199; DB 2; Length 971;
Best Local Similarity 20.6%; Pred. No. 0.013;
Matches 133; Conservative 79; Mismatches 227; Indels 206; Gaps 20;

QY 9 IASNSSSSTSRSDVD-----STTATATPTPPP-----TFDD 41
Db 38 VASENPEAGRYSMDEVGYQGVSVILLVEGPPSHAGTITTVYEDSQPTLNDFLGAMTDD 97
QY 42 YKTAQQTAYDTIF-----TSTSLADIQAALVSLQDAVNTNIKDTAAATDEBETAABEWTK 95

```
Db      98 VRPEALRRFELMVREAVARNASAVAGNTAAAKKSADASTSARE-----AATHATD 147
Qy      96 NADAVKVGQAQITELAKYASDNOAILDSLGKLTSLFDLQALLQSVANNKAAELIKEMQD 155
Db      148 AADSAR--AASTSAQAASSAQAASSAGTAST-----KATEASKSAAAESSKSAATSAAGTSETNAVSO-Q 218
Qy      156 NPVVPGKTPALQAQSLVDQTDATATQIEKDGNNAIRDYFAGQNASGAVENAKSNNSISNID 215
Db      179 -----KATEASKSAAAESSKSAATSAAGTSETNAVSO-Q 216
Qy      216 SAKAAIATAKTQIAEAAOKKPPDPSPILOEAMQVIOAEKDLKNIKPADGSDVNPNGTTVGG 275
Db      217 SAATSASTATTYKASEAASARDASASKEA-----AKSETSAASSASSASS 263
Qy      276 SKOOGSSIGSTRVSMMLDDAENETASILMSGFROMIHMENTENPDQSAQAQELAAQAPAA 335
Db      264 ATAAGNSAKAAKTS-----ETNAKSETAAEQSAS 293
Qy      336 KAAGDDSAALADAQAALGKAGQOQOQILNALGOIASAAVVSAGV-----P 385
Db      294 AAAGSKTAAALSA-----SAASTSAGQASASATAAGKSAESAASSASTATTKAGEATEQ 347
Qy      386 PAAASSIGSSVKQL---YKTSKSTGSDYKTOISAGYDAYKINDAYGRANDATR--DVI 440
Db      348 ASAAAASSASAATSETNAKASSETSAESSKTAASASSASSASSASSASASASKDEATROASAA 407
Qy      441 NNVTSPALTRSPRA-----RTEARGPEKTDQALARVIGNSRSLGD 482
Db      408 KSSATTASTKATEAGSATAAQAQSKSTAESAATRAETAARAEADIASAVALDEDASTTKKG 467
Qy      483 VYSOVSAQSVMQITQSNPQ-----ANNEEIROKLTSAVTKPPQPGYPPVOLS--D 532
Db      468 IVQLSSATNSTSESIAATPKAVKAAAYDLANGKYTAQDATT-----QKG--IVQLSNATN 520
Qy      533 STOKFPIA-----KLESLFAGSRTAAE-----IKALSFETNS 564
Db      521 STSEMLAATPKSVKAAAYDLANGKYTAQDATTAAQKGIVOLSSATNS 565

RESULT 9
C85693
probable membrane protein of prophage CP-933X Z1918 [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: C85693
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85693
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STO>
A:Cross-references: GB:AB005174; NID:G12514847; PIDN:AAG56007.1; GSPDB:GN00145; UWGP:Z19
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1918

Query Match      7.0%; Score 199; DB 2; Length 973;
Best Local Similarity 20.6%; Pred. No. 0.013;
Matches 133; Conservative 79; Mismatches 227; Indels 206; Gaps 20;

Qy      9 IASNSSSSTSRADVD-----STATAATPPPP-----TFDD 41
Db      40 VASNPDPDAGRISMDVEXGQSVILLVEGFPFPHAGTIVVEDSQPGTLNDFLGAMTDD 99
Qy      42 YKTAQATYDTIF-----TSTSLADIAQALVSLQDAVTNIKDTAATDEETAIAAEWETK 95
Db      100 VRPEALRRFELMVREAVARNASAVAGNTAAAKKSADASTSARE-----AATHATD 149
Qy      96 NADAVKVGQAQITELAKYASDNOAILDSLGKLTSLFDLQALLQSVANNKAAELIKEMQD 155
```

```
Db      150 AADSAR--AASTSAQAASSAQAASSAGTAST----- 180
Qy      156 NPVVPGKTPALQAQSLVDQTDATATQIEKDGNNAIRDYFAGQNASGAVENAKSNNSISNID 215
Db      181 -----KATEASKSAAAESSKSAATSAAGTSETNAVSO-Q 218
Qy      216 SAKAAIATAKTQIAEAAOKKPPDPSPILOEAMQVIOAEKDLKNIKPADGSDVNPNGTTVGG 275
Db      219 SAATSASTATTYKASEAASARDASASKEA-----AKSETSAASSASSASS 265
Qy      276 SKOOGSSIGSTRVSMMLDDAENETASILMSGFROMIHMENTENPDQSAQAQELAAQAPAA 335
Db      266 ATAAGNSAKAAKTS-----ETNAKSETAAEQSAS 295
Qy      336 KAAGDDSAALADAQAALGKAGQOQOQILNALGOIASAAVVSAGV-----P 385
Db      296 AAAGSKTAAALSA-----SAASTSAGQASASATAAGKSAESAASSASTATTKAGEATEQ 349
Qy      386 PAAASSIGSSVKQL---YKTSKSTGSDYKTOISAGYDAYKINDAYGRANDATR--DVI 440
Db      350 ASAAAASSASAATSETNAKASSETSAESSKTAASASSASSASSASSASASKDEATROASAA 409
Qy      441 NNVTSPALTRSPRA-----RTEARGPEKTDQALARVIGNSRSLGD 482
Db      410 KSSATTASTKATEAGSATAAQAQSKSTAESAATRAETAARAEADIASAVALDEDASTTKKG 469
Qy      483 VYSOVSAQSVMQITQSNPQ-----ANNEEIROKLTSAVTKPPQPGYPPVOLS--D 532
Db      470 IVQLSSATNSTSESIAATPKAVKAAAYDLANGKYTAQDATT-----QKG--IVQLSNATN 522
Qy      533 STOKFPIA-----KLESLFAGSRTAAE-----IKALSFETNS 564
Db      523 STSEMLAATPKSVKAAAYDLANGKYTAQDATTAAQKGIVOLSSATNS 567

RESULT 10
F90073
hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: F90073
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90073
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-2271 <KOR>
A:Cross-references: GB:BA000018; PID:G13702612; PIDN:BA843752.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2447

Query Match      6.9%; Score 195.5; DB 2; Length 2271;
Best Local Similarity 18.6%; Pred. No. 0.06;
Matches 110; Conservative 132; Mismatches 240; Indels 109; Gaps 20;

Qy      2 SLAD--KLGIASNSSSSTSRADVDSTTATPTPPPTFDYKTAQATYDTIFTSTS 58
Db      1236 SLSDTSSEGSTSTJUNSTJSGSASISTSTSISEST-----STFKSESVSTJUSMTSTS 1290
Qy      59 LADIQAALVSLQDAVTNIK--DTAATDEETAIAAEWETKADAVKVGQAQIT--ELAKYASDN 116
Db      1291 LSNSTSLSTSLSDSTSDSKDSLSTSMISDSI--STSKSDSISTSTSLSGSTSESEDS 1348
Qy      117 QAILDSLGKLTSLFDLQALLQSVANNKAAELIKEMQDNPVVPGKTPALQAQSLVDQTD 176
Db      1349 TSSSEKSDSTSMIS--SMSQSTS-----GSTSTSTSLSDSTS 1386
Qy      177 TATQIEKDGNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAAIATATAQIAEAAOKKFP 236
```


Job time : 20.9063 secs

```
RESULT 15
C82206
methyl-accepting chemotaxis protein Vcl1403 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82206
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.I.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406:477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: C82206
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-641 <HEI>
A:Cross-references: GB:AE004218; GB:AE003852; NID:g9655881; PIDN:AAF94560.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: Vcl1403
A:Map position: 1

Query Match          6.3%; Score 179.5; DB 2; Length 641;
Best Local Similarity 20.2%; Pred. No. 0.074;
Matches 125; Conservative 110; Mismatches 202; Indels 181; Gaps 28;

QY      14 SSSSTSRSDVDSTTATPTPTPTFDYKTKQAOT-----AYDTFTSTSLADIQAALV 67
Db      9 TSASSSPSPNHSANAPKTEA---DDLRFQARTLARKQQAERVASAT--AELLAGLE 63

QY      68 SLQDAVNIKDTAATDEETAIAEWE---TKNADAV--KVGAGITELAKYASDNQAILDS 122
Db      64 ESSAMTQIRSSM---EQIAYGAEBSSTATKQSETAVSQMNFINEQATVATQSKILTK 120

QY      123 LGLKTSFDLQALLOALQSVAANNKAAELLKENODNPVPGKTPAQAQSLVDQDTATATCIE 182
Db      121 LEQ-----DIVDA-----NKINMVSNVQ-----TSSERQNSVKKRMIE 155

QY      183 KDGNAIRDAYFAGONASGAVENAKSNNSISNI---DSAKAAIATAKTQIAEAQKPPDS 238
Db      156 LSEQA-----AKINDAVQVHIADQTNLLNALNAIEAGRAGKHGKGF 198

QY      239 PILQEAQOMVIQ-AEKDLKNTK-----ADGSDVPNPGTTVGGSK-----QQGSSIGS-- 285
Db      199 AVVADTVRTLAEKAEKNAATESLIKDTISNGAEMVSVNGVSKSSQKAESEVQKGHWVNKQL 258

QY      286 --IRVSM--LUDDAE-----NETASILMSGFRQMIHMFENTENPDQAQQLAAQA 332
Db      259 TLIRNEMVDLVKDEVLKKADEMSSAAVIAKGSSEV-----SQAAEQSAAACE 308

QY      333 RAAKAAGDSDAAALADAQAALGAALGAGQOQOQILNALGQIAGAAVVSAGVPPAAASSI 392
Db      309 ESLK-----SLDQOQIALDGAVTAAQSLDELTDDEL-RTSTDIVKSSSEVAAAEEL 358

QY      393 GSSVKQLYKTSKTSQDYKTOISAGYDAYKINDAYGBARDATRDVNNVST--PALTR 450
Db      359 SAGIEEINRSSNEI-MGALNQISSG-----ACHMAKSVETGITSLTQ 399

QY      451 SVPRARTEARGPEKTDQALAEVIGS---NSRTLGDVYKSOVSALQSVMOITOS----- 499
Db      400 IEQARLCKERAESLKLKACEKMLGIEENKTTVDEM-----ILATTESTKAATENL 450

QY      500 NPQANNEIRQKLTSAV-----TKPPQFGYPYVQLSNDSTQKFI 538
Db      451 NEMANIERISQIDKIVDGISNVSIQTAMLVANGVAEAAAGAYGKGFVVSTD----- 504

QY      539 AKLESFPAEGSRTAABIK 556
Db      505 --IQNLANDAAENAEQIK 520
```

Search completed: March 24, 2004, 06:00:08

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:24:24 ; Search time 50.4852 Seconds
(without alignments)
3262.843 Million cell updates/sec

Title: US-10-608-559-4

Perfect score: 2850

Sequence: 1 MSLADKLGIASSNSSSTSR.....SLFIQQVLVNGISYGLQ 583

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2844	99.8	583	AAV71955	Aay71955 5'-trunca
2	2844	99.8	651	AAV71954	Aay71954 Chlamydia
3	2844	99.8	651	AAU38899	Aau38899 C. pneumo
4	2844	99.8	651	ABE94277	ABE94277 Chlamydia
5	2844	99.8	651	ABE90532	ABE90532 Chlamydia
6	2844	99.8	651	ABE90532	ABE90532 Chlamydia
7	1870	65.6	490	AAV71957	Aay71957 C. pneumo
8	1866	65.5	452	AAV71956	Aay71956 3'-trunca
9	1844	64.7	478	AAV35358	Aay35358 Chlamydia
10	995	34.9	212	AAV35357	Aay35357 Chlamydia
11	843	29.6	647	ABE94283	ABE94283 Chlamydia
12	843	29.6	647	ADD43866	ADD43866 Chlamydia
13	825.5	29.0	660	AAU38921	Aau38921 C. tracho
14	825.5	29.0	660	ADD42726	ADD42726 Chlamydia
15	666.5	23.4	361	AAE13695	AAE13695 Chlamydia
16	666.5	23.4	361	AAE13695	AAE13695 Chlamydia
17	666.5	23.4	361	AAE13695	AAE13695 Chlamydia
18	599.5	21.0	331	ABE94234	ABE94234 Protein e
19	227.5	8.0	350	AAV37572	Aay37572 Chlamydia
20	214.5	7.5	1178	AAV37571	Aay37571 Chlamydia
21	211.5	7.4	1463	ABJ26182	ABJ26182 Aspergill
22	211.5	7.4	1831	AAE20110	AAE20110 lactobaci
23	211.5	7.4	1870	ABU43109	ABU43109 Protein e
24	211	7.4	603	ABU44080	ABU44080 Protein e
25	207	7.3	1822	AAE27745	AAE27745 Extracell

ALIGNMENTS

RESULT 1

AAV71955
ID AAY71955 standard; protein; 583 AA.

XX AC AAY71955;

XX DT 26-MAR-2001 (first entry)

XX DE 5'-truncated Chlamydia pneumoniae 76 kDa protein.

XX 76 kDa protein; bactericidal; diagnosis; prevention; pneumonia;
XX upper respiratory tract disease; bronchitis; sinusitis;
XX acute respiratory disease; cough; sore throat; hoarseness; fever;
XX vaccine; immunisation; treatment; truncation mutant; mutein.

XX OS Chlamydia pneumoniae.

XX OS Synthetic.

XX PH Location/Qualifiers

XX FT Misc-difference 497 /note= "Encoded by ACT"

XX PN WO200066739-A2.

XX PD 09-NOV-2000.

XX PF 03-MAY-2000; 2000WO-CA000511.

XX PR 03-MAY-1999; 99US-0132270P.

XX PR 30-JUN-1999; 99US-0141276P.

XX PA (AVET) AVENTIS PASTEUR LTD.

XX PI Murdin AD, Oomen RP, Wang J, Dunn P;

XX WPI; 2000-687542/67.

XX N-PSDB; AAD02064.

XX Nucleic acids encoding a 76 kDa protein from Chlamydia pneumoniae, useful for vaccinating against Chlamydia infections.

XX Claim 16b; Page 100-102; 112pp; English.

XX The present sequence is 5'-truncated Chlamydia pneumoniae 76 kDa protein.
XX C. pneumoniae 76 kDa protein is used in the diagnosis, prevention and
XX treatment of C. pneumoniae infections (e.g. pneumonia, upper respiratory
XX tract disease, bronchitis, sinusitis and acute respiratory disease such
XX as cough, sore throat, hoarseness, fever; and abnormal chest sounds on

Handwritten signature

QY 361 AGQQGILNALGQIASAAVVSAGV 384
 Db 429 ACQQGILNALGQIASAAVVSAGV 452

RESULT 9
 AAY35358
 ID AAY35358 standard; protein; 478 AA.
 XX AAY35358;
 AC AAY35358;
 XX 17-OCT-2003 (revised)
 DT 13-SEP-1999 (first entry)
 XX Chlamydia pneumoniae surface exposed polypeptide.
 DE
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 KW neutralising epitope.
 XX Chlamydia pneumoniae.
 OS Chlamydia pneumoniae.
 XX WO9927105-A2.
 PN 03-JUN-1999.
 XX 20-NOV-1998; 98WO-IB001890.
 XX 21-NOV-1997; 97FR-00014673.
 PR 04-NOV-1998; 98US-0107078P.
 XX (GEST) GENSET.
 PA Griffais R;
 XX WPI; 1999-357842/30.
 DR Genome sequence of Chlamydia pneumoniae.
 XX Page 1156; Disclosure; 1912pp; English.
 PS AAY34584-Y35879 represent the proteins encoded by all the open reading
 XX frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis,
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)
 XX Sequence 478 AA;
 SQ

Query Match 64.7%; Score 1844; DB 2; Length 478;
 Best Local Similarity 100.0%; Pred. No. 1.1e-111; Indels 0; Gaps 0;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLADKLGASNSSTSRSDVSTTATPTPPPTFDDYKTAQAYDTFTSTSLA 60
 Db 71 MSLADKLGASNSSTSRSDVSTTATPTPPPTFDDYKTAQAYDTFTSTSLA 130

QY 61 DIQALVSLQDVTNFKOTATDEETAIAAETWKNADAVKVGQITELAKYASDNQAIL 120
 Db 131 DIQALVSLQDVTNFKOTATDEETAIAAETWKNADAVKVGQITELAKYASDNQAIL 190

QY 121 DSLGKLTSDLLQALLOSVAANNKAAELLKEMQNPVPGKTPAIAQSLVDQDTATATQ 180
 Db 191 DSLGKLTSDLLQALLOSVAANNKAAELLKEMQNPVPGKTPAIAQSLVDQDTATATQ 250

QY 181 FEKGNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAATATAKTQTAEAKKFPDPSPI 240

Db 251 FEKGNNAIRDYFAGQNASGAVENAKSNNSISNIDSAKAATATAKTQTAEAKKFPDPSPI 310
 QY 241 LOEAEQMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMMLDDAENETA 300
 Db 311 LOEAEQMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMMLDDAENETA 370
 QY 301 SILMSGFRQMIHMENTENPDSQAAQELAAQARAKAAGDDSSAAAALADAQKALEAALGK 360
 Db 371 SILMSGFRQMIHMENTENPDSQAAQELAAQARAKAAGDDSSAAAALADAQKALEAALGK 430

QY 361 AGQQGILNALGQIASAAV 379
 Db 431 AGQQGILNALGQIASAAV 449

RESULT 10
 AAY35357
 ID AAY35357 standard; protein; 212 AA.
 XX AAY35357;
 AC AAY35357;
 XX 17-OCT-2003 (revised)
 DT 13-SEP-1999 (first entry)
 XX Chlamydia pneumoniae transmembrane protein sequence.
 DE
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 KW neutralising epitope.
 XX Chlamydia pneumoniae.
 OS Chlamydia pneumoniae.
 XX WO9927105-A2.
 PN 03-JUN-1999.
 XX 20-NOV-1998; 98WO-IB001890.
 XX 21-NOV-1997; 97FR-00014673.
 PR 04-NOV-1998; 98US-0107078P.
 XX (GEST) GENSET.
 PA Griffais R;
 XX WPI; 1999-357842/30.
 DR Genome sequence of Chlamydia pneumoniae.
 XX Page 1155; Disclosure; 1912pp; English.
 PS AAY34584-Y35879 represent the proteins encoded by all the open reading
 XX frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis,
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)
 XX Sequence 212 AA;
 SQ

Query Match 34.9%; Score 995; DB 2; Length 212;
 Best Local Similarity 99.0%; Pred. No. 6.5e-57;
 Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 380 VSAGVPPAAASSSTSSVKQLYKPSKSTGSDYKTKTQISAGYDAYKSINDAYGRANDTRDV 439
 Db 9 VSAGVPPAAASSSTSSVKQLYKPSKSTGSDYKTKTQISAGYDAYKSINDAYGRANDTRDV 68

```

QY 440 INNVSTPALTRSPRARTARGPEKTDQALARVISGNSRILGDVYSQVSALQSVMQITQS 499
Db 69 INNVSTPALTRSPRARTARGPEKTDQALARVISGNSRILGDVYSQVSALQSVMQITQS 128
QY 500 NPOANNEIRKLTSAVTKPQPGYPPVOLSNDSOKFIKLSLFAEGSRTAETIKALS 559
Db 129 NPOANNEIRKLTSAVTKPQPGYPPVOLSNDSOKFIKLSLFAEGSRTAETIKALS 188
QY 560 FEINSLFIQVVLNIGSLYGYLQ 583
Db 189 FEINSLFIQVVLNIGSLYGYLQ 212

RESULT 11
ID ABB94283 standard; protein; 647 AA.
XX AC ABB94283;
XX DT 05-JUN-2002 (first entry)
XX Chlamydia trachomatis protein sequence SEQ ID NO:436.
XX Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
XX antibody; immunostimulant; immune response;
XX Chlamydia-specific T-cell response.
XX Chlamydia trachomatis.
XX OS WO200208267-A2.
XX PN 31-JAN-2002.
XX PD 20-JUL-2001; 2001WO-US023121.
XX PF 20-JUL-2000; 2000US-00620412.
XX PR 23-APR-2001; 2001US-00841132.
XX PA (CORI-) CORIXA CORP.
XX Fling SP, Skeiky YAW, Probst P, Bhatia A;
XX WPI; 2002-179901/23.
XX Novel compositions comprising Chlamydia Cap1 protein and its use in the
XX treatment of Chlamydia infection.
XX Claim 34; Page 391-393; 537pp; English.
XX The present invention describes compositions comprising a Chlamydia Cap1
XX protein and methods for the diagnosis and therapy of Chlamydia infection.
XX Chlamydia DNA and protein sequences from the present invention can have
XX antibacterial and immunostimulant activities, and can be used in
XX vaccines. Compounds from the present invention can be used for eliciting
XX an immune response, specifically stimulating a Chlamydia-specific T-cell
XX response or inhibiting the development of a Chlamydia infection in an
XX animal. Methods from the present invention can be used: for detecting the
XX presence of Chlamydia in a patient; to stimulate and/or expand T cells
XX specific for a Chlamydia protein; and for treatment of a Chlamydia
XX infection. ABL92394 to ABL92709 and ABB94036 to ABB94374 represent
XX sequences used in the exemplification of the present invention
XX Sequence 647 AA;
XX SQ

Query Match 29.6%; Score 843; DB 5; Length 647;
Best Local Similarity 35.5%; Pred. No. 2.3e-46;
Matches 213; Conservative 113; Mismatches 243; Indels 31; Gaps 14;

QY 1 MSLADKLGIASSSSSSRS-ADVDSITTATPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT 59
Db 60 ISLDD--AILNKNSSPDSUSQLBAS--STVTVAARDNEAKSNFTAKSGLENATTL-117
QY 60 ADIQAAVLSQLDAVTNIKDTAATDEE-TAIAAEWETKNADAVKVGQITELAKYASDQA 118

```

```

Db 118 ABYETKWLADLMAALQDMLERLAKQAEVTRIKEALQEKQ-----EVIDKLNQLVKLEKQNOT 173
QY 119 ILDSLGLKLSFDLLQALLOSANNKKAELKEMODNFPVPGKTPAIAQSLVDQTDATA 178
Db 174 LKELTITTSADQIPAINSOLEINKNSADQIIKLEGGNI---SYEAVLTNAGEVTKAS 230
QY 179 TQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAAKAIATATAKQIAEAQKK---F 235
Db 231 EAGIKLQALQSLVDAGDOSQAAVLQAQCNNSPDNIAATKCLIDAAETKVNELKQSHHTGL 290
QY 236 PDSPILOBAEQMVIOAEKOLKNIKPADGSDVPNPGTTVGSQKQSSISGI-----RV 288
Db 291 TDSPLVKKAEQISOAKDOIQEIKPS-GSDIPTVGPS-GSAASAGAVGALKSSNNSGRI 348
QY 289 SMLDDAENETASILMSGFRQIMHMENTENPDSQAAQQLAAQARAAK---AAGDDSA 345
Db 349 SLILDDVDNEMAAIAMCGFRSMIEQFNVNPNPATAKELQAVEAQLTAMSDQLVGADGELPA 408
QY 346 ALADAQKALEALGKAGQOQGIILNALGQIASAAVVSAGVPPPAASSIGSVKQLYKT--S 403
Db 409 EIQAIKDALAQAL-KQPTDGLATAMQVAFAAAKVGGSGAGTAGTVQMNVKOLYKTAFS 467
QY 404 KSTGSDYKTIQISAGYDAYKSINDAYGEABNDATRDVINNVSTPALTRSPRARTARG-P 462
Db 468 STSSSYAAALSDGYSAYKTLNLSYSEKRS-GVQSAISQITANPALSRVSRSIGESQGRS 526
QY 463 EKTDQALARVISGNSRITLGDVYSQVSALQSVMQITQSNNPQANNEIRKLTSAVTKPPQF 522
Db 527 ADASQRAAETIVRDSQTLGDVYSRLQVLDSLMSTIVSNPQVNCQEEIMQKLITASISKAPOF 586
QY 523 GYFYVQLSNDSTOKFIKLSLFAEGSRTAETIKALSFEINSLFIQVVLNIGSLYGYL 582
Db 587 GYFVQVQVADSLQKFAAQLEREFDVGDERSLAESRENAFRKQPAFIQVVLNIASLPSGYL 646

RESULT 12
ADD43866
ID ADD43866 standard; protein; 647 AA.
XX AC ADD43866;
XX DT 15-JAN-2004 (first entry)
XX Chlamydia trachomatis immunogenic protein, SEQ ID NO 161.
XX immunogenic; infection; Chlamydia trachomatis; immune; vaccine;
XX gene therapy; antibacterial.
XX Chlamydia trachomatis.
XX OS WO2003049762-A2.
XX PN 19-JUN-2003.
XX PD 12-DEC-2002; 2002WO-IB005761.
XX PR 12-DEC-2001; 2001GB-00029732.
XX PR 06-AUG-2002; 2002GB-00018233.
XX PR 14-AUG-2002; 2002GB-00018924.
XX (CHIR-) CHIRON SPA.
XX Grandi G, Ratti G;
XX WPI; 2003-532882/50.
XX N-PSDB; ADD43867.
XX New immunogenic composition having a protein or encoding nucleic acid,
XX useful for diagnosing, preventing and/or treating Chlamydia trachomatis
XX infection.
XX Claim 6; SEQ ID NO 161; 164pp; English.
XX PS

```

XX The invention relates to a novel immunogenic composition comprising a
 CC protein or nucleic acid, and an adjuvant, where the protein or nucleic
 CC acid comprises any of 131 fully defined amino acid or nucleotide
 CC sequences given in the specification, or has 50% or greater sequence
 CC identity to it, or their fragments. The protein and/or nucleic acid of
 CC the immunogenic composition is useful in the manufacture of a medicament
 CC for the treatment or prevention of infection due to Chlamydia
 CC trachomatis. The infection is treated or prevented by the medicament
 CC eliciting an immune response which is specific to a C. trachomatis
 CC elementary body, or for neutralising C. trachomatis elementary bodies,
 CC hence the immunogenic composition can be used in creating a vaccine. The
 CC immunogenic compositions can also be used for the diagnosis of C.
 CC trachomatis infection. The nucleic acids of the immunogenic compositions
 CC can be used to treat disorders by gene therapy. The immunogenic
 CC compositions have antibacterial activity. This sequence represents one of
 CC the 131 C. trachomatis proteins with immunogenic properties of the
 CC invention.
 XX
 SQ Sequence 647 AA;

Query Match 29.6%; Score 843; DB 7; Length 647;
 Best Local Similarity 35.5%; Pred. No. 2.3e-46;
 Matches 213; Conservative 113; Mismatches 243; Indels 31; Gaps 14;
 QY 1 MSLADKLGIASSNSSSSSTSR-ADVDSTTATPTPPPTTDDYKTOAQAYDTFTTSL 59
 Db 60 ISLRD--ALLNKNSSPTDLSQLEASTSTVTRVAARDYNEAKSNFTAKSGLENATLL 117
 QY 60 ADIQALVSLQDAVNIKDTATDEE-TALAAEWETKNADAVKGAQITELAKYASNDQA 118
 Db 118 AYEYTKWADLMAALQDMERLAKQAEVTRIKALQEKQ-----EVIDKLNQLVKLEKQNT 173
 QY 119 ILDSLGKLTSPDLLQALLQSVANNKKAELLKEMQDNVVPKTPATAQSLVDOTDATA 178
 Db 174 LKEITLTTDSADQIPAINQSLQLEINKNADQIIKDLQGNII---SYEAVLTNAGEVIKASS 230
 QY 179 TQIEKGNNAIRDVAFAGQVAGSVAENAKSNNSISNIDSAKALATAKTQIAEAKKK--F 235
 Db 231 EAGIKLQALQSLVDAGQSQAAVLQAOQNSPDNIATKLLIDAEYKVNELKQEHGL 290
 QY 236 PDSPIQAEQVQAEKDLKNKPKADGSDVNPPTTGVGSKQKQSSIGSI-----RV 288
 Db 291 TDSPTVKKAEQISQAQKDIQIEIKPS-GSDIPTVGPS-GSAASAGSAGVALKSSNSGRI 348
 QY 289 SMLDDAENETASILMSGFRQMIHMTENTPDSQAQQLAAQARA--AAGDSSAA 345
 Db 349 SILLDDVDNEMAAIAMQGRFVIEQFNVPNNPATAKELQAMEAQLTAMSDQLVGADGELPA 408
 QY 346 ALADAQKALEAALGKAGQOQILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKT--S 403
 Db 409 EIQAIKDALAQAL-KQPSDGLATAMGQVAFAPAAAKVGGSGAGTAGTVQMNVKQLYXTAFS 467
 QY 404 KTSQSDYKTOISAGDYAKSINDAYGRARNDAIRVNNVSTPALTRVPRARTEARG-P 462
 Db 468 STSSSYAAALSDGYSAYKTLNLSYSESR-GVQSAISQTANPALSRVSRSGIESQGRS 526
 QY 463 EKTQALARVLSGNSRTLDGVYVSQVSLQSVNQLTQSNPQANNEIROKLTSAVTKPQF 522
 Db 527 ADASQRAAETIVRDSQTLGDVYSRLQVLDLSMTIVSNPQVNOEIMOKLTASISKAPQF 586
 QY 523 GYPVYQLSNDSTQKFIKLESIFARSGRTAAEIKALSFEFNSLFTQQVLVNIIGSLYGYL 582
 Db 587 GYPVQNSADSLQKFAAQLEREFVDGERSLAESRENAFRKQPAFTQQVLVNIASIFSGYL 646
 RESULT 13
 AAU38921
 ID AAU38921 standard; protein; 660 AA.
 XX
 AC AAU38921;
 XX
 DT 11-SEP-2003 (revised)

DT 16-JAN-2002 (first entry)
 XX
 DE C. trachomatis CT875 protein.
 XX
 KW Chlamydia; sexually transmitted disease; PID; antibacterial;
 KW pelvic inflammatory disease; antigen; trachoma; gynecological;
 KW acute respiratory tract infection; atherosclerosis; male infertility;
 KW coronary heart disease.
 XX
 OS Chlamydia trachomatis; serovar E.
 XX
 PN WO200181379-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 23-APR-2001; 2001WO-US013081.
 XX
 PR 21-APR-2000; 2000US-0198853P.
 PR 20-JUL-2000; 2000US-0219752P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Bhatia A, Probat P, Stromberg EJ;
 XX
 DR N-PSDB; AAS56995.
 XX
 DR WPI; 2001-616771/71.
 XX
 DR N-PSDB; AAS56995.
 XX
 PT New polynucleotide for treating Chlamydia infections encodes a
 PT polynucleotides containing an immunogenic portion of a Chlamydia antigen.
 XX
 FS Claim 3; Page 205-207; 208pp; English.
 XX
 CC The invention relates to isolated polynucleotide encoding at least a
 CC partial Chlamydia protein which is an antigenic fragment, or the
 CC complements, fragments, homologues and variants, and antibodies raised
 CC against the antigenic proteins (or fragments). The nucleic acids,
 CC proteins and antibodies are used to diagnose and treat Chlamydia
 CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
 CC disease (PID), acute respiratory tract infection, trachoma, and in the
 CC atherosclerosis and coronary heart disease) in a patient, and in the
 CC treatment of male infertility. The compounds of the invention are also
 CC useful for detecting the presence of Chlamydia in a patient, and
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC The present sequence represents a Chlamydia antigen. (Updated on 11-SEP-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 660 AA;
 Query Match 29.0%; Score 825.5; DB 4; Length 660;
 Best Local Similarity 34.8%; Pred. No. 3.2e-45;
 Matches 212; Conservative 117; Mismatches 236; Indels 45; Gaps 15;
 QY 1 MSLADKLGIASSNSSSSSTSR-ADVDSTTATPTPPPTTDDYKTOAQAYDTFTTSL 59
 Db 67 ISLRD--ALLNKNSSPTDLSQLEASTSTVTRVAARDYDEAKSNFTAKSGLENATLL 124
 QY 60 ADIQALVSLQDAVNIKDTAATD-----EET-----ATAEWETKNADAVKGAQITE 108
 Db 125 AYEYTKWADLMAALQDMERLANSDFNNHTEEVNNIKALEAQKDTID-----KLNK 176
 QY 109 LAKYASNDQAALDSGLKLTSPDLLQALLQSVANNKKAELLKEMQDNVVPKTPAIAQ 168
 Db 177 LVTLQNKQSLTEVLYKTTDSADQIPAINQSLQLEINKNADQIIKDLQGNII---SYEAVLT 233
 QY 169 SLVDQDTATQTEKQGNARIDAFAGQVAGSVAENAKSNNSISNIDSAKALATAKTQI 228
 Db 234 NAGEVIKASSEAGIKLQALQSLVDAGDQSAAVLQAOQNSPDNIATKELIDAETKV 293
 QY 229 AEAOKK---FPDSPILQAEQVQAEKDLKNKPKADGSDVNPPTTGVGSKQKQSSIGS 285
 Db 294 NELQKHTGTGTDSPLVKKABEQISQAQKDIQIEIKPS-GSDIPTVGPS-GSAASAGSAGA 351
 QY 286 I-----RVSMELDDAENETASILMSGFRQMIHMTENTPDSQAQQLAAQARA-- 336

KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;
KW coronary heart disease; antibacterial.
XX
XX Chlamydia sp.
XX
XX WO200034483-A2.
XX
XX 15-JUN-2000.
XX
XX 08-DEC-1999; 99WO-US029012.
XX
XX 08-DEC-1998; 98US-00208277.
XX 08-APR-1999; 99US-00288594.
XX 01-OCT-1999; 99US-00410568.
XX 22-OCT-1999; 99US-00426571.
XX
XX (CORI-) CORIXA CORP.
XX
XX Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
XX WPI; 2000-431303/37.
XX
XX Isolated polypeptide for diagnosis and treatment of Chlamydia infection
XX comprises immunogenic portion of Chlamydia antigen, which comprises amino
XX acid sequence encoded by polynucleotide sequence.
XX
XX Disclosure; Page 251-253; 256pp; English.
XX
XX The present invention relates to new nucleic acid sequences and the
XX proteins encoded by the nucleic acid sequences. The encoded proteins
XX comprise an immunogenic portion of a Chlamydia antigen. The encoded
XX proteins are useful for the serodiagnosis and treatment of Chlamydia
XX infection. Chlamydiae are intracellular bacterial pathogens that are
XX responsible for a wide variety of human infections. C. trachomatis
XX infection is one of the most common sexually transmitted diseases and can
XX lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
XX and infertility. Trachoma due to ocular infection with C. trachomatis is
XX the leading cause of preventable blindness worldwide. C. pneumoniae is a
XX major cause of acute respiratory tract infections in humans and is also
XX thought to play a role in the pathogenesis of atherosclerosis and
XX coronary heart disease. The present sequence is a protein isolated in the
XX present invention
SQ Sequence 361 AA;

Query Match 23.4%; Score 666.5; DB 3; Length 361;
Best Local Similarity 42.1%; Pred. No. 3.2e-35;
Matches 153; Conservative 62; Mismatches 131; Indels 17; Gaps 8;
QY 233 KFPDPPILOAEQVIOAEKLNKIPADGSDVDPNGTGVGGSKQGGSSIGSI----- 286
DB 2 QEIAADPLVKAEEQINQAQDQITPTFS-GLDIPVIGPS-GSAASAGSAGALKSSNNS 59
QY 287 -RVSMLLDDAENETASILMSGFROHMENTENPDPSQAAQELAAQARAAK---AAGDDS 342
DB 60 GRISLDDVDVNEMAIANQGFNSIEQFNVPNNPATAKELQAMEAQLTAMSDQLVGADGE 119
QY 343 AAALADAKALAEALGKAGQQQGLNALGQITASAAVVSAGVPPAAASSIGSSVFQLYKT 402
DB 120 LPAETIQAIKDALAQAL-KQPSADGLATANGQVAFAAKVGGSGAGTACTVQMNVKQLYKT 178
QY 403 --SKTSGDYKTQISAGDYDAYKINDAYGRANDATRDVINNVSTPALTRSVPRARTEAR 460
DB 179 AFSSTSSSSYAAALSDGYSAYKTLNLSXSESS-GVQSAISQTANPALSRVSRSRGIESQ 237
QY 461 G-PEKTDQALRVISGNSRTGLDGVYSOVSALOSVMOITCSNPQANNEIROKLTSAVTKP 519
DB 238 GRSADASQEAETIVRDSQTLGDVTSRQVLDLSMTIVSNPQANQEEIMQKLTASISKA 297
QY 520 PQFGYPYVQLSNDSTQKFIKLESIFAEGSRTAABIKALSFPETNSLFTQQVLVNIIGSLYS 579

Db 298 PQFGYPVQNSVDSLQKFAAQLEREFVDGERSLAESQENAFKQKQPAFTQQVLVNIASLFS 357
QY 580 GYL 582
Db 358 GYL 360
Search completed: March 23, 2004, 19:27:39
Job time : 52.4852 secs

QY 121 DSLGKLTSPDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQDATATQ 180
 DB 121 DSLGKLTSPDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQDATATQ 180
 QY 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 DB 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 QY 241 LQAEQMWIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOQSSIGSIRVSMMLDDAENETA 300
 DB 241 LQAEQMWIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOQSSIGSIRVSMMLDDAENETA 300
 QY 301 SILMSGFRQMHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 360
 DB 301 SILMSGFRQMHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 360
 QY 361 AGQOQGILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKTOISAGYDA 420
 DB 361 AGQOQGILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKTOISAGYDA 420
 QY 421 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 480
 DB 421 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 480
 QY 481 GDVYSQVSALQSVQWQITQSNPQANNEIRKLTSAVTKPPQFGYVYVQVLSNDSTQKPIAK 540
 DB 481 GDVYSQVSALQSVQWQITQSNPQANNEIRKLTSAVTKPPQFGYVYVQVLSNDSTQKPIAK 540
 QY 541 LESLFAEGSRPTAAEIKALSPFTNSLFTQOVLVNTGSLYSGYLQ 583
 DB 541 LESLFAEGSRPTAAEIKALSPFTNSLFTQOVLVNTGSLYSGYLQ 583

RESULT 2
 US-09-841-132-405
 ; Sequence 405, Application US/09841132
 ; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841,132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 405
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-841-132-405

Query Match 99.8%; Score 2844; DB 9; Length 651;
 Best Local Similarity 99.8%; Pred. No. 2e-204;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSLADKLGIASSSSSSRSADVSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 60
 DB 69 MSLADKLGIASSSSSSRSADVSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 128
 QY 61 DIQAALVSLQDVTNIKDTAATDEETAAAEWETKNADAVKGAQITELAKYASDNQAIL 120
 DB 129 DIQAALVSLQDVTNIKDTAATDEETAAAEWETKNADAVKGAQITELAKYASDNQAIL 188
 QY 121 DSLGKLTSPDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQDATATQ 180
 DB 189 DSLGKLTSPDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQDATATQ 248
 QY 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 DB 249 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 308

QY 241 LQAEQMWIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOQSSIGSIRVSMMLDDAENETA 300
 DB 309 LQAEQMWIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOQSSIGSIRVSMMLDDAENETA 368
 QY 301 SILMSGFRQMHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 360
 DB 369 SILMSGFRQMHMFNTENPDSCAAQOELAAQARAAGDSDSAAAALADAQKALEAALGK 428
 QY 361 AGQOQGILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKTOISAGYDA 420
 DB 429 AGQOQGILNALGQIASAAVVSAGVPPAAASSIGSVKQLYKTSKSTGSDYKTOISAGYDA 488
 QY 421 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 480
 DB 489 YKSDINAYGRANDATRDVINNVSTPALTRSPVPRTEARGPEKTDQALARVSGNSRTL 548
 QY 481 GDVYSQVSALQSVQWQITQSNPQANNEIRKLTSAVTKPPQFGYVYVQVLSNDSTQKPIAK 540
 DB 549 GDVYSQVSALQSVQWQITQSNPQANNEIRKLTSAVTKPPQFGYVYVQVLSNDSTQKPIAK 608
 QY 541 LESLFAEGSRPTAAEIKALSPFTNSLFTQOVLVNTGSLYSGYLQ 583
 DB 609 LESLFAEGSRPTAAEIKALSPFTNSLFTQOVLVNTGSLYSGYLQ 651

RESULT 3
 US-09-564-479-2
 ; Sequence 2, Application US/09564479
 ; Publication No. US20030095973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: COWEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; CURRENT FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR APPLICATION NUMBER: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-564-479-2

Query Match 99.8%; Score 2844; DB 10; Length 651;
 Best Local Similarity 99.8%; Pred. No. 2e-204;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSLADKLGIASSSSSSRSADVSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 60
 DB 69 MSLADKLGIASSSSSSRSADVSTTATPTPPPTFDDYKTOAQYATDTTSTSLA 128
 QY 61 DIQAALVSLQDVTNIKDTAATDEETAAAEWETKNADAVKGAQITELAKYASDNQAIL 120
 DB 129 DIQAALVSLQDVTNIKDTAATDEETAAAEWETKNADAVKGAQITELAKYASDNQAIL 188
 QY 121 DSLGKLTSPDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQDATATQ 180
 DB 189 DSLGKLTSPDLLQALLOSANNKAAELLKEMQDNVVPCKTPAIAQSLVDQDATATQ 248
 QY 181 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 240
 DB 249 IEKDGNAIRDAYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFFDPSPI 308
 QY 241 LQAEQMWIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOQSSIGSIRVSMMLDDAENETA 300

Db 309 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 368
 QY 301 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 360
 Db 369 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 428
 QY 361 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSTGSDYKTIQISAGYDA 420
 Db 429 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSTGSDYKTIQISAGYDA 488
 QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 480
 Db 489 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 548
 QY 481 GDVYSQVSALQSVQMCIQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 540
 Db 549 GDVYSQVSALQSVQMCIQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 608
 QY 541 LESLFAEGSRTAAETKALSFTFNSLFIQOVLVNIQSLYSGYLQ 583
 Db 609 LESLFAEGSRTAAETKALSFTFNSLFIQOVLVNIQSLYSGYLQ 651

RESULT 4

US-09-841-260-78
 ; Sequence 78, Application US/09841260
 ; Publication No. US20030175700A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; APPLICANT: Stromberg, Erika Jean
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
 ; FILE REFERENCE: 210121.515
 ; CURRENT APPLICATION NUMBER: US/09/841-260
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 140
 ; SEQ ID NO 78
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-841-260-78

Query Match 99.8%; Score 2844; DB 10; Length 651;
 Best Local Similarity 99.8%; Pred. No. 2e-204;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSLADKLGIIASSNSSSSRSADVDSTTATATPTPPPTFDDYKTAQATYDTFTSTSLA 60
 Db 69 MSLADKLGIIASSNSSSSRSADVDSTTATATPTPPPTFDDYKTAQATYDTFTSTSLA 128
 QY 61 DIOAALVSLQDAVTNICKOTAATDEETAAEAETKNADAVKVGQAQITELAKYASDNQAIL 120
 Db 129 DIOAALVSLQDAVTNICKOTAATDEETAAEAETKNADAVKVGQAQITELAKYASDNQAIL 188
 QY 121 DSLGKLTSPDLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 180
 Db 189 DSLGKLTSPDLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 248
 QY 181 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQKPPDSP 240
 Db 249 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQKPPDSP 308
 QY 241 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 300
 Db 309 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 368
 QY 301 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 360
 Db 429 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQKPPDSP 308
 QY 241 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 300
 Db 309 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 368
 QY 301 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 360
 Db 369 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 428
 QY 421 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 420
 Db 489 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 428
 QY 361 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSTGSDYKTIQISAGYDA 420

Db 429 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSTGSDYKTIQISAGYDA 488
 QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 480
 Db 489 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 548
 QY 481 GDVYSQVSALQSVQMCIQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 540
 Db 549 GDVYSQVSALQSVQMCIQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 608
 QY 541 LESLFAEGSRTAAETKALSFTFNSLFIQOVLVNIQSLYSGYLQ 583
 Db 609 LESLFAEGSRTAAETKALSFTFNSLFIQOVLVNIQSLYSGYLQ 651

RESULT 5

US-10-007-693-78
 ; Sequence 78, Application US/10007693
 ; Publication No. US20020146776A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; FILE REFERENCE: 210121/515C2
 ; CURRENT APPLICATION NUMBER: US/10/007-693
 ; CURRENT FILING DATE: 2001-12-05
 ; NUMBER OF SEQ ID NOS: 157
 ; SEQ ID NO 78
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-10-007-693-78

Query Match 99.8%; Score 2844; DB 13; Length 651;
 Best Local Similarity 99.8%; Pred. No. 2e-204;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSLADKLGIIASSNSSSSRSADVDSTTATATPTPPPTFDDYKTAQATYDTFTSTSLA 60
 Db 69 MSLADKLGIIASSNSSSSRSADVDSTTATATPTPPPTFDDYKTAQATYDTFTSTSLA 128
 QY 61 DIOAALVSLQDAVTNICKOTAATDEETAAEAETKNADAVKVGQAQITELAKYASDNQAIL 120
 Db 129 DIOAALVSLQDAVTNICKOTAATDEETAAEAETKNADAVKVGQAQITELAKYASDNQAIL 188
 QY 121 DSLGKLTSPDLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 180
 Db 189 DSLGKLTSPDLLQALLOSVAANNKAAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 248
 QY 181 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQKPPDSP 240
 Db 249 IEKQGNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATATKQIAEAQKPPDSP 308
 QY 241 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 300
 Db 309 LQEAQWVIOAEKDLKNIKPAGSDVPNPPTTVGSKQGGSSIGSVKQYKTSKSTGSDYKTIQISAGYDA 368
 QY 301 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 360
 Db 369 SILMSGFQMIHMFNTENPDSQAQOELAAQAAKAAAGDSDAAAALADAQKALEAALGK 428
 QY 361 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSTGSDYKTIQISAGYDA 420
 Db 429 AGQOQGIILNALGQIASAAVVSAGVPPAAASSIGSVKQYKTSKSTGSDYKTIQISAGYDA 488
 QY 421 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 480
 Db 489 YKSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVISGNSRTL 548
 QY 481 GDVYSQVSALQSVQMCIQSNPQANNEIRQKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 540

Db 549 GDVYSQVSALQSVWQIIQSNPQANNEIRKQLTSVTKPQPGYPPVQLSNDSTQKFIK 608
 Qy 541 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLNIGSLYGYLQ 583
 Db 609 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLNIGSLYGYLQ 651

RESULT 6

US-10-312-273-13
 ; Sequence 13, Application US/10312273
 ; Publication No. US20040005667A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SpA
 ; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
 ; FILE REFERENCE: P025035W0
 ; CURRENT APPLICATION NUMBER: US/10/312,273
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: 0016363.4
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 0017047.2
 ; PRIOR FILING DATE: 2000-07-11
 ; PRIOR APPLICATION NUMBER: 0017983.8
 ; PRIOR FILING DATE: 2000-07-21
 ; PRIOR APPLICATION NUMBER: 0019368.0
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: 0020440.4
 ; PRIOR FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: 0022583.9
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 0027549.5
 ; PRIOR FILING DATE: 2000-11-10
 ; PRIOR APPLICATION NUMBER: 0031706.5
 ; PRIOR FILING DATE: 2000-12-22
 ; NUMBER OF SEQ ID NOS: 664
 ; SOFTWARE: SeqWin99, version 1.02
 ; SEQ ID NO 13
 ; LENGTH: 651
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-10-312-273-13

Query Match 99.8%; Score 2844; DB 15; Length 651;
 Best Local Similarity 99.8%; Pred. No. 2e-204;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSLADKLGIASSNSSSTSRGADVDSTTATPTPTPTPTFDYKTAQATYDTTFTSTSLA 60
 Db 69 MSLADKLGIASSNSSSTSRGADVDSTTATPTPTPTPTFDYKTAQATYDTTFTSTSLA 128
 Qy 61 DIQAALVSLQDAVTVNIKDTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 120
 Db 129 DIQAALVSLQDAVTVNIKDTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 188
 Qy 121 DSLGKLTSLFLLQALLOSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 180
 Db 189 DSLGKLTSLFLLQALLOSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 248
 Qy 192 IEKQGNATRDYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 240
 Db 249 IEKQGNATRDYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 308
 Qy 241 LQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 300
 Db 309 LQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 368
 Qy 301 SILMSGFRQMIHMTENTENPDSQAQELAAQARAARAAAGDDSAALADAKALEALGK 360
 Db 369 SILMSGFRQMIHMTENTENPDSQAQELAAQARAARAAAGDDSAALADAKALEALGK 428
 Qy 361 AGCQGGIILNALGQIASAAVWSAGVPPAAASSIGSSVKQLYKTSKSTGDKYKTQISAGYDA 420
 Db 429 AGCQGGIILNALGQIASAAVWSAGVPPAAASSIGSSVKQLYKTSKSTGDKYKTQISAGYDA 488

Qy 421 YKSINDAYGEARNDATRDVNNVSTALTRESVPRATEARGPEKTDQALARVSGNSRTL 480
 Db 489 YKSINDAYGEARNDATRDVNNVSTALTRESVPRATEARGPEKTDQALARVSGNSRTL 548
 Qy 481 GDVYSQVSALQSVWQIIQSNPQANNEIRKQLTSVTKPQPGYPPVQLSNDSTQKFIK 540
 Db 549 GDVYSQVSALQSVWQIIQSNPQANNEIRKQLTSVTKPQPGYPPVQLSNDSTQKFIK 608
 Qy 541 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLNIGSLYGYLQ 583
 Db 609 LESLFAEGSRSTAABEIKALSFETNSLFIQQVLNIGSLYGYLQ 651

RESULT 7

US-09-564-479-8
 ; Sequence 8, Application US/09564479
 ; Publication No. US20030095973A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MURDIN, ANDREW D.
 ; APPLICANT: COHEN, RAYMOND P.
 ; APPLICANT: WANG, JOE
 ; APPLICANT: DUNN, PAMELA
 ; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
 ; TITLE OF INVENTION: USES THEREOF
 ; FILE REFERENCE: 032931/0230
 ; CURRENT APPLICATION NUMBER: US/09/564,479
 ; CURRENT FILING DATE: 2000-05-03
 ; PRIOR APPLICATION NUMBER: 60/132,270
 ; PRIOR FILING DATE: 1999-05-03
 ; PRIOR APPLICATION NUMBER: 60/141,276
 ; PRIOR FILING DATE: 1999-06-30
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 490
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-564-479-8

Query Match 65.6%; Score 1870; DB 10; Length 490;
 Best Local Similarity 99.7%; Pred. No. 1.3e-131;
 Matches 385; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSLADKLGIASSNSSSTSRGADVDSTTATPTPTPTPTFDYKTAQATYDTTFTSTSLA 60
 Db 69 MSLADKLGIASSNSSSTSRGADVDSTTATPTPTPTPTFDYKTAQATYDTTFTSTSLA 128
 Qy 61 DIQAALVSLQDAVTVNIKDTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 120
 Db 129 DIQAALVSLQDAVTVNIKDTAATDEETAIAAEWETKNADAVKVGACITELAKYASDNQAIL 188
 Qy 121 DSLGKLTSLFLLQALLOSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 180
 Db 189 DSLGKLTSLFLLQALLOSVANNKKAELLKEMQDNVPVPGKTPAIAQSLVDQTDATATQ 248
 Qy 181 IEKQGNATRDYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 240
 Db 249 IEKQGNATRDYFAGQNASGAVENAKSNNSISNIDSAKAAATATAKTQIAEAQKFPDPSPI 308
 Qy 241 LQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 300
 Db 309 LQEAQWVIQAEKDLKNIKPADGSDVPNPGTTVGGSKQGGSSIGSIRVSMLLDDAENETA 368
 Qy 301 SILMSGFRQMIHMTENTENPDSQAQELAAQARAARAAAGDDSAALADAKALEALGK 360
 Db 369 SILMSGFRQMIHMTENTENPDSQAQELAAQARAARAAAGDDSAALADAKALEALGK 428
 Qy 361 AGCQGGIILNALGQIASAAVWSAGVPP 386
 Db 429 AGCQGGIILNALGQIASAAVWSAGVLP 454

RESULT 8

US-09-564-479-6
; Sequence 6, Application US/09564479
; Publication No. US20030095973A1
; GENERAL INFORMATION:
; APPLICANT: MURDIN, ANDREW D.
; APPLICANT: COWEN, RAYMOND P.
; APPLICANT: WANG, JOE
; APPLICANT: DUNN, FAMELA
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: 032931/0230
; CURRENT APPLICATION NUMBER: US/09/564,479
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,270
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: 60/141,276
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-564-479-6

Query Match
Best Local Similarity 65.5%; Score 1866; DB 10; Length 452;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 60
Db 69 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 128
Qy 61 DIQALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGQITELAKYASNQAIL 120
Db 129 DIQALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGQITELAKYASNQAIL 188
Qy 121 DSLGKLTSDLLOALLOSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 180
Db 189 DSLGKLTSDLLOALLOSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 248
Qy 181 IEKGNNAIRDAYFAGQNASGAVENAKSNNSINIDSAKAAIATATQIAEAQKPPDSP 240
Db 249 IEKGNNAIRDAYFAGQNASGAVENAKSNNSINIDSAKAAIATATQIAEAQKPPDSP 308
Qy 241 LOEAQOMVIQAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMMLDDAENETA 300
Db 309 LOEAQOMVIQAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMMLDDAENETA 368
Qy 301 SILMSGFRQMIHMENTENPDSQAQOELAAQARAARAAAGDDSAALADAQKALEAALGK 360
Db 369 SILMSGFRQMIHMENTENPDSQAQOELAAQARAARAAAGDDSAALADAQKALEAALGK 428
Qy 361 AGQQGILNALGQIASAAVVSAGV 384
Db 429 AGQQGILNALGQIASAAVVSAGV 452

RESULT 9
US-10-289-762-776
; Sequence 776, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 776
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae

US-10-289-762-776
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-776

Query Match
Best Local Similarity 64.7%; Score 1844; DB 15; Length 478;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 60
Db 71 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDDYKTOAQAYDTFTSTSLA 130
Qy 61 DIQALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGQITELAKYASNQAIL 120
Db 131 DIQALVSLQDAVNTIKDTAATDEETAAAEWETKNADAVKVGQITELAKYASNQAIL 190
Qy 121 DSLGKLTSDLLOALLOSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 180
Db 191 DSLGKLTSDLLOALLOSVANNKAAELLKEMQNPVPGTTPAIAQSLVDQTDATATQ 250
Qy 181 IEKGNNAIRDAYFAGQNASGAVENAKSNNSINIDSAKAAIATATQIAEAQKPPDSP 240
Db 251 IEKGNNAIRDAYFAGQNASGAVENAKSNNSINIDSAKAAIATATQIAEAQKPPDSP 310
Qy 241 LOEAQOMVIQAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMMLDDAENETA 300
Db 311 LOEAQOMVIQAEKDLKNIKPADGSDVPNPPTTVGGSKQGGSSIGSIRVSMMLDDAENETA 370
Qy 301 SILMSGFRQMIHMENTENPDSQAQOELAAQARAARAAAGDDSAALADAQKALEAALGK 360
Db 371 SILMSGFRQMIHMENTENPDSQAQOELAAQARAARAAAGDDSAALADAQKALEAALGK 430
Qy 361 AGQQGILNALGQIASAAV 379
Db 431 AGQQGILNALGQIASAAV 449

RESULT 10
US-10-289-762-775
; Sequence 775, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; FILE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 775
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-775

Query Match
Best Local Similarity 34.9%; Score 995; DB 15; Length 212;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 380 VSAGVPPAAASSIGSSVKOLYKTSKSTGSDYKTOISAGYDAYKSNIDAYGRANDATRDV 439
Db 9 VSAGVPPAAASSIGSSVNLKTSKSTGSDYKTOISAGYDAYKSNIDAYGRANDATRDV 68
Qy 440 INNVTPTALTSVPRATEARPEKTDQALARVIGSNRSLTGDVYSQVSAQSVNQITQS 499
Db 69 INNVTPTALTSVPRATEARPEKTDQALARVIGSNRSLTGDVYSQVSAQSVNQITQS 128
Qy 500 NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKFIKLSFABGSRTAABIKALS 559
Db 129 NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKFIKLSFABGSRTAABIKALS 188
Qy 560 FETNSLFTQQVLNIGLSYGLQ 583


```
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 139
; LENGTH: 660
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-007-693-139
```

```
Query Match      29.0%; Score 825.5; DB 13; Length 660;
Best Local Similarity 34.8%; Pred. No. 3.6e-53;
Matches 212; Conservative 117; Mismatches 236; Indels 45; Gaps 15;

QY 1 MSLADKLGAGSSSSSSRS--ADVSTTATPTPPPTTDDYKTAQATYDTFTSTSL 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 67 ISLRD--AILKNSSPTDSLSQLEASTSTSTVTRVAAKDYDEAKGNFDTAKSGLENAKTL 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 ADIQAAVLSQDAVINIKDTATD-----EET-----AIAAEWETKNADAVKVGCAQITE 108
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 ABYETKMDLMAALQMERLANSDFSNHTEEVNNIKALEAQKDTID-----KLNK 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 109 LAKVASDNOAILDSLGLKTSFLLQALLQSVAANNKAAELLKEMQDNPVVPKTPALAAQ 168
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 LVTLQNKSLTEVLKTTDSADQIIPAINSOLEINKNSADQIIKDLERQNI---SYEAVLT 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 169 SLVDTQDTATQIEKGNRAIRDAYFAGQNAAGAVENAKNSNISNIDSAKAIAIATKQI 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 NAGEVIKASSEAGIKLQALQSIDVADGQSAALVQAQONNSPDNIAATKELIDAEATKV 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 229 ABAQKK---FPDSPILOAEQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGS 285
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 NELKQHTGLTDSPLVKKAEQISQAQKDIQEIKPS--GSDIPIVGPS--GSAASAGSAGA 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 I-----RVSMLLDDAENETASILMSGFQMIHMFNTENPDSQAAQQLAAQARAANK-- 336
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 LKSSNNSGRISLLDDVDNEMAALQGRSMIEQFNVPNNPATAKELQAMEAQLTAMSDQ 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 337 -AAGDDSAALADAKALEALGKAGQOQOQILNALGQIASAAVVSAGVPPAAASIGSS 395
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 412 LVGADGELPAETQAIKDALAQAL-KQPSADGLATAMGQVAFAAAKVGGSGAGTAGTVQMN 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 VKOLYKT--SKTSGDYKTOISAGVDAYSINDAYGRANDATRDVINNVSTPALTRSP 453
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 VKOLYKTAFSTSSSSSYAALSDGYSAYKTLNSLYSESR--GVQSAISQTANPALSRVS 529
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 454 RARTEARG--PEKTDQALARVISGNRTLGDVYSQVSALQSNVMOITQSNPQANNEIROKL 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 530 RSGIERSQGRSADASQRAAETIVRDSQTLGDVYSRLQVLDLSMTTIVSNPQANQEEIMQKL 589
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 513 TSAVTKPPFGYPYVOLSNDSTQKFIKLESIFAEGRSRTAAEIKALSFTNSLFTQOVLV 572
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 590 TASISKAPOFGYPVQNSADSLQKFAAQLEFVDGERSLAESQENAFKQPAFIQQOVLV 649
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 573 NIGSLYSGL 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 650 NIASLFSGYL 659
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 14
US-09-841-132-299
; Sequence 299, Application US/09841132
; Patent No. US2002006184BA1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
```

```
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 299
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-841-132-299
```

```
Query Match      23.4%; Score 666.5; DB 9; Length 361;
Best Local Similarity 42.1%; Pred. No. 1.2e-41;
Matches 153; Conservative 62; Mismatches 131; Indels 17; Gaps 8;

QY 233 KKFDPSPILOAEQOMVIOAEKDLKNIKPADGSDVPNPGTTVGGSKQOGSSIGSI----- 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 QEIDADSPLVKKAEQIQQAQDIIQTITPS--GLDIPVIGPS--GSAASAGSAGALKSSNNS 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 287 -RVSMLLDDAENETASILMSGFQMIHMFNTENPDSQAAQQLAAQARAANK--AAGDDS 342
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 GRISILLDDVDNEMAALQGRSMIEQFNVPNNPATAKELQAMEAQLTAMSDQLVGADGE 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 343 AAALADAKALEALGKAGQOQOQILNALGQIASAAVVSAGVPPAAASIGSSVKLYKT 402
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 LPATQIAIKDALAQAL-KQPSADGLATAMGQVAFAAAKVGGSGAGTAGTVQMNVKLYKT 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 403 --SKTSGDYKTOISAGVDAYSINDAYGRANDATRDVINNVSTPALTRSPVPRARTEAR 460
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 AFSSTSSSYAALSDGYSAYKTLNSLYSESR--GVQSAISQTANPALSRVSRSIGSEQ 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 461 G-PEKTDQALARVISGNRTLGDVYSQVSALQSNVMOITQSNPQANNEIROKLTSVATKP 519
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 GRSADASQRAAETIVRDSQTLGDVYSRLQVLDLSMTTIVSNPQANQEEIMQKLTAISKA 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 520 PQFGYPYVOLSNDSTQKFIKLESIFAEGRSRTAAEIKALSFTNSLFTQOVLVNIAGSLYS 579
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 PQFGYPVQNSVDSLQKFAAQLEFVDGERSLAESQENAFKQPAFIQQOVLVNIASLFS 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 580 GYL 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 GYL 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 15
US-10-128-714-8240
; Sequence 8240, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroehkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8240
; LENGTH: 1178
; TYPE: PRT
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 13:39:25 ; Search time 10.7195 Seconds
(without alignments)
2831.945 Million cell updates/sec

Title: US-10-608-559-4

Perfect score: 2850

Sequence: 1 MSLADKLGIASSSSSTSR.....SLFTQQVLNIGSLYSGYILQ 583

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	6.3	2541	1 TLN1_HUMAN	Q9V490 homo sapien
2	176	6.2	8545	1 ANCI1_CAEEL	Q9N4M4 caenorhabdi
3	174	6.1	2541	1 TLN1_MOUSE	P26039 mus musculus
4	173.5	6.1	1140	1 YN36_YEAST	Q04893 saccharomyc
5	172	6.0	1306	1 MS22_YEAST	P32334 saccharomyc
6	171.5	6.0	774	1 STE1_LAMBDA	P03764 bacterioph
7	171	6.0	1120	1 STFR_ECOLI	P76072 escherichia
8	167	5.9	641	1 HRPK_PSESY	P41501 pseudomonas
9	165.5	5.8	2542	1 TLN2_HUMAN	Q9V4G6 homo sapien
10	163.5	5.7	881	1 PRV3_YEAST	P47033 saccharomyc
11	163.5	5.7	2090	1 N214_HUMAN	P35658 homo sapien
12	163	5.7	3644	1 MINT_MOUSE	Q62504 mus musculu
13	161.5	5.7	2492	1 TALAI_DICDI	P54633 dictyosteli
14	160	5.6	1161	1 DANA_YEAST	P47179 saccharomyc
15	158	5.5	1562	1 SPAP_STRMU	P22504 streptococc
16	158	5.5	1565	1 PAC_STRMU	P11657 streptococ
17	158	5.5	1690	1 C193_DROME	Q9VJES drosophila
18	157.5	5.5	810	1 HTR4_HALSA	Q48317 halobacteri
19	156.5	5.5	810	1 HTR4_HALNI	Q9NP84 halobacteri
20	153	5.4	490	1 SE05_ECOLI	P47500 escherichia
21	152.5	5.4	1528	1 SPAA_STRDO	P21979 streptococ
22	150.5	5.3	758	1 IMWT_HUMAN	P16891 homo sapien
23	150	5.3	626	1 CE1B_ECOLI	P04479 escherichia
24	149.5	5.2	763	1 HTR2_HALNI	Q9HP81 halobacteri
25	149.5	5.2	764	1 HTR2_HALSA	P71410 halobacteri
26	149.5	5.2	1902	1 P2P_LACLC	P15293 lactococcus
27	149	5.2	516	1 P54_ENTFC	P13692 enterococcu
28	149	5.2	757	1 IMWT_MOUSE	Q8CAQ8 mus musculu
29	148.5	5.2	507	1 RRPD_CDUO	P06940 canine dist
30	148	5.2	1238	1 SBCC_RHOCA	Q68032 rhodobacter
31	147.5	5.2	1011	1 SC4A_RICAF	Q9AJ83 rickettsia
32	147.5	5.2	1902	1 P1P_LACLC	P16271 lactococcus
33	146.5	5.1	962	1 IF2_NEIMA	Q9JTB5 neisseria m

RESULT 1

ID	TLN1_HUMAN	STANDARD;	PRT;	2541 AA.
AC	Q9V490; Q9NZQ2; Q9UHH8;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Talin 1.			
GN	TLN1 OR TLN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1] SEQUENCE FROM N.A.			
RP	MAO L., Fan Y.H.;			
RA	"Complete cDNA sequence of human talin.";			
RT	Submitted (JUL-1998) to the EMBL/GenBank/DBAJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
EX	MEDLINE=20079169; PubMed=10610730;			
RA	Ben-Yosef T., Francomano C.A.;			
RT	"Characterization of the human talin (TLN) gene: genomic structure, chromosomal localization, and expression pattern.";			
RT	Genomics 62:316-319(1999).			
CC	FUNCTION: Probably involved in connections of major cytoskeletal structures to the plasma membrane. High molecular weight cytoskeletal protein concentrated at regions of cell-substratum contact and, in lymphocytes, at cell-cell contacts (By similarity).			
CC	SUBUNIT: Binds with high affinity to vinculin and with low affinity to integrins (By similarity).			
CC	SIMILARITY: Contains 1 FERM domain.			
CC	SIMILARITY: Contains 1 I/LWEQ domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AF078828; AAD3152.1; --			
DR	EMBL; AF177198; AAF23322.1; --			
DR	EMBL; AF178534; AAF27330.1; --			
DR	EMBL; AF178081; AAF27330.1; JOINED.			
DR	Gene; HGNC:11845; TLN1.			
DR	MIM; 186745; --			
DR	GO; GO:0005925; C:focal adhesion; NAS.			
DR	GO; GO:0005200; P:structural constituent of cytoskeleton; NAS.			
DR	GO; GO:0006928; P:cell motility; NAS.			
DR	GO; GO:0007016; P:cytoskeletal anchoring; NAS.			
DR	InterPro; IPR000299; Band 4.1.			
DR	InterPro; IPR002558; ILWEQ.			
DR	Pfam; PF00373; Band_41; 1.			

ALIGNMENTS

34	146	5.1	625	1	VP35_BPAPS	Q9CLR3 bacterioph
35	146	5.1	1093	1	TMF1_HUMAN	P82094 homo sapien
36	145.5	5.1	707	1	YU9C_YEAST	P47166 saccharomyc
37	145.5	5.1	1022	1	SCA4_RICCN	Q52658 rickettsia
38	145.5	5.1	1080	1	MRTB_MOUSE	P59759 mus musculu
39	145	5.1	3178	1	Y889_CAEEL	Q09624 caenorhabdi
40	145	5.1	3259	1	GOB1_HUMAN	Q14789 homo sapien
41	144.5	5.1	1509	1	MYSN_ACACA	P05659 acanthamoeb
42	144	5.1	1325	1	Y309_MYCPN	P75334 mycoplasma
43	144	5.1	1609	1	LMG1_HUMAN	P11047 homo sapien
44	143.5	5.0	991	1	SCA4_RICSI	Q9AJ77 rickettsia
45	143.5	5.0	5171	1	BPEA_HUMAN	O94833 homo sapien

DR	Pfam; PF01608; I LWEQ; 1.	
DR	ProDom; PD011820; ILWEQ; 1.	
DR	SMART; SM00295; B41; 1.	
DR	SMART; SM00307; ILWEQ; 1.	
DR	PROSITE; PS00660; FERM 1; 1.	
DR	PROSITE; PS00661; FERM 2; 1.	
DR	PROSITE; PS50057; FERM 3; 1.	
DR	PROSITE; PS50945; I LWEQ; 1.	
KW	Structural protein; Cytoskeleton.	
FT	DOMAIN 86 403	FERM
FT	DOMAIN 2340 2533	I/LWEQ
FT	CONFLICT 824 824	G -> R (IN REF. 2; AAF27330).
FT	CONFLICT 1227 1227	S -> L (IN REF. 1).
FT	CONFLICT 1549 1549	P -> A (IN REF. 2; AAF27330).
FT	CONFLICT 1604 1604	Q -> K (IN REF. 2; AAF27330).
FT	CONFLICT 1701 1701	E -> Q (IN REF. 2; AAF27330).
FT	CONFLICT 1718 1718	H -> N (IN REF. 2; AAF27330).
FT	CONFLICT 1966 1966	A -> R (IN REF. 1).
FT	CONFLICT 2256 2256	MISsing (IN REF. 2; AAF27330).
FT	CONFLICT 2541 2541	AA; 269717 MW; 47ED9B22BCF47296 CRC64;
SQ	SEQUENCE	
Query Match 6.3%; Score 180; DB 1; Length 2541;		
Best Local Similarity 21.6%; Pred. No. 0.27;		
Matches 148; Conservative 97; Mismatches 256; Indels 184; Gaps 32;		
QY	49 AYDTFTST-----SLADI-----QAALV--SLQDVTNINIKTAADEE-----T 86	
DB	801 ATDTILVTENIFSGMGDAGEVMGQARILAQATSDLVNAIKADAEGBESDLENSRKLLSAA 860	
QY	87 AIAEWEKTNADAVK-----VGAQITELAKYASDNOAILDLSGLKLTSD 130	
DB	861 KILADATAKMVEAAKAAHDPDSEEQOQRUREAEAGLEGMTATNAQAQNAIKKLVQRLEHA 920	
QY	131 LLQAA--LLOSVAANNKAAELLKE-----MODNFVVPKTPAIQASLVQDQDVAATQ 180	
DB	921 AKQAAASATQTIAAQAHAASPKASAGPOPLLQSKAVAFQIPLLQGV-----RGSQ 974	
QY	181 IEKQGNIRDAYFAGN-----ASGAVENAKSNNSISNISAKAAIATATQI 228	
DB	975 AQPDPSPAQLALIASQSFLQGGKMWAAAKASVPTIQDQASAMQLSQCAKNLGLTALAE 1034	
QY	229 AEAQKKPDS--PILOEAEQWVIO-AEKDLAKIIPA--DGSVDNPGTGVGSKQ--GSS 282	
DB	1035 RTAAQKAQEAQCPLEMDSALSVMQLEKDLQEVKAAARDGKLLPLPGETKECTQDGLNS 1094	
QY	283 IGS1--RVSMLLDAB--NE-----TASILMSGFRQMIHMFN-----TENPDSQAQOEL 328	
DB	1095 TKAVSSAIFQALLGEVAQGNENYAGTARDVAGGLRLSLAQAAAGVAAULTSDPAVQAIVLD 1154	
QY	329 AAQA-----RAAKAA--GDDSAALADAAQKALEAALGKAGQ-----QQGLNALLG 372	
DB	1155 ASDVLDKASSLIEEAKAAGHFGDPESQOQLAQAVAKVQTALNRCVSCLPQORDVDNALR 1214	
QY	373 QI--ASAAVVSAGVPPA-----AASSTGSSVKOLYKTSKST----- 406	
DB	1215 AVGDASKLLSDSLPPSTGTQEQAGRLNEAAGLNQATELVQASRTGTPQDLAPASGRF 1274	
QY	407 GSDYKTOISAGYDAYKSIINDYGRANDATR--DVINN-----VSPALTRSVPRARTEAR 460	
DB	1275 GQDFSTFLEAGV-----EMAGAPSQEDRAQVVSNLKGISMSSSLKLLAAKALSTDP 1327	
QY	461 GREKTDQ--ALARVIGNSRNLGVDYSQ-----VSLQSVMQITQSNPQANNE-- 506	
DB	1328 APNLKSQLAAAAARAVTDSINQLITCTQQAQPGQKCDNALRELETVRELLENFPVQIPNDM 1387	
QY	507 -----ETROKLTSAVT-----KPPQFGYPVQVLSNDSTQKFAKLESFAE 547	
DB	1388 SYFGCLDSVWNSKVLGEANTGISQNAKNGNLPERG-----DAISTASKALCGTE 1438	
QY	548 GSRTAAEIKALSFETNSLFIQQLV 572	
DB	1439 AAAQAAYLVGVS--DPNSAQGQQLV 1462	

RESULT 2

ID	ANCL CABEL	STANDARD;	PRT; 8545 AA.
AC	Q9N4M4; O61841; O61842;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Nuclear anchorage protein 1 (Anchorage 1 protein) (Nesprin homolog).		
GN	ANC-1 OR ZK973.6 OR T03A1.3/T03A1.4.		
OS	Caenorhabditis elegans.		
OC	Rhabditida; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Beloderinae; Caenorhabditis.		
OX	NCHI_taxid=6239;		
RN	[1]		
RP	SEQUENCE OF 1-1751 FROM N.A., FUNCTION, CHARACTERIZATION, AND		
RP	INTERACTION WITH F-ACTIN AND UNC-84.		
RC	STRAIN=Bristol N2;		
RX	MEDLINE=22264052; PubMed=12169658;		
RA	Starr D.A., Han M.;		
RT	"Role of ANC-1 in tethering nuclei to the actin cytoskeleton.";		
RL	Science 298:406-409(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	Gattung S., Goela D., Broy M.;		
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	REVISIONS.		
RA	Waterston R.;		
RL	Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.		
RN	[4]		
RP	SEQUENCE OF 689-1243 FROM N.A.		
RX	MEDLINE=22296983; PubMed=12408964;		
RA	Zhang Q., Ragnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;		
RT	"The nesprins are giant actin-binding proteins, orthologous to		
RT	Drosophila melanogaster muscle protein MSP-300.";		
RL	Genomics 80:473-481(2002).		
CC	-!- FUNCTION: Plays a central role in nuclear and mitochondrial		
CC	anchoring. Probably connects nuclei to the cytoskeleton by		
CC	interacting with unc-84 at the nuclear envelope and with F-actin		
CC	in the cytoplasm, creating a bridge across the nuclear envelope		
CC	between the cytoskeleton and the nucleus.		
CC	-!- SUBUNIT: Interacts with F-actin via its N-terminal domain.		
CC	Interacts indirectly with unc-84.		
CC	-!- SUBCELLULAR LOCATION: Type IV membrane protein (potential). The		
CC	largest part of the protein is cytoplasmic, while its C-terminal		
CC	part is associated either with the nuclear envelope, most probably		
CC	the outer nuclear membrane, or with mitochondrial membrane.		
CC	-!- TISSUE SPECIFICITY: Ubiquitously expressed in all postembryonic		
CC	cells.		
CC	-!- DEVELOPMENTAL STAGE: First expressed in L1 larvae, and thereafter.		
CC	-!- DOMAIN: The large coiled coil domains are composed of 6 nearly		
CC	exact repeats of 903 residues. The last repeat is partial. These		
CC	repeats are conserved in Hawai (CB4856), Australia (AB4) and		
CC	Germany (RC301) strains. The length of the repeat may be		
CC	maintained because of a selective advantage to keep the protein		
CC	large and allow a single molecule to extend more than 0,5		
CC	micrometers.		
CC	-!- DOMAIN: The Klarsicht domain, which contains a potential		
CC	transmembrane domain, is essential for the nuclear envelope		
CC	targeting.		
CC	-!- SIMILARITY: Belongs to the nesprin family.		
CC	-!- SIMILARITY: Contains 1 actin-binding domain.		
CC	-!- SIMILARITY: Contains 2 calponin-homology (CH) domains.		
CC	-!- SIMILARITY: Contains 1 Klarsicht domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)		

ED COIL (POTENTIAL).

[illegible]


```
GN YMR317W OR YM9924.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93 (1997).
CC -!- DOMAIN: Contains many Ser/Thr-rich domain and repeats.
CC -!- SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; 254141; CA90835.1; -.
DR GenOnline; 142998; -.
DR SGD; S0004936; YMR317W.
KW Hypothetical protein; Repeat.
SQ SEQUENCE 1140 AA; 113070 MW; 0153EBCA24FE5427 CRC64;

Query Match
Best Local Similarity 6.1%; Score 173.5; DB 1; Length 1140;
Matches 115; Conservative 127; Mismatches 277; Indels 163; Gaps 15;

QY 11 SSSSSSSSRADVDSTTATATPPP-----PTFDYKQTQATVDTTFTS 58
DB 23 STSTSTTTPTTSTTSTSTTKVTTSPETIVSSSTLVSSVPEFTSSSLSDTTIASILS 82
QY 59 LADIQAAVLSDAVTNKINDTAADETAIAA-----EWETKNADAVKVG 103
DB 83 SESLVISFSLYSTSSDISSTSVNDVESSTGSPNSYALSSTNAQLSSITETDSISS 142
QY 104 AQITELEAYASDQAILDGLKLTGFDLLQALLOSANNKAAEL-----LKMQDN 156
DB 143 ALQTSSPQTSSNG-----GGSEPLEGKSVLETTASSDPTTAVTSTFTLLIDVSS 196
QY 157 PVVPGKTPAI-----AQSIVDQTDATATQIEKQNAIRDYFA-- 194
DB 197 PKISSGGNAVTSVGTSTDSKXEVFSSSTSDVSSLLSSTSSPASSISETLPSSITLIT 256
QY 195 -----GQNASGAVENAKNSNISIDSAKAAIATAKTQIAEA-----QKFPDPP 239
DB 257 SSPVSSEAPSATSSVSSEASSTSSVSSEAPLATSSVVSSEAPSSSTSVVSSEAPST 316
QY 240 ILQEAQEMVIQAEKDLKNIKPADGDVDPNGFTTVGSKQOQSSISIRVSMLLDAENET 299
DB 317 SSSVSESEI-SSTTSSVSSEAPLATSSVSSEAPSTSSVSSEI-SSTTSSVSSEAPLAT 376
QY 300 ASILMSGFRQMTHTMENTNPDSQAQQLAAQAAKAGDDPSAAALADAQKALEAALG 359
DB 377 SSVVSS-----EAPSTSS--SVSSEAPSTSSVSSEAPSTSSVSSEI-SST 423
QY 360 KAGQQQGIINALGQIASAAVWSAGVPPA-----AASIGSS 395
DB 424 KS-----SVMSSEVSATSSLSVSEAPSAISLSSRLFPSSKNTSVTSLTVAEASVTS 479
QY 396 VKQLYKTSKTSQDYKTIQISAGYDAYKINDAYGRANDATRDVINNVSTPALTSVPFA 455
DB 480 LRPSETTLAS-NSIISSLSLTGYNSTVS-----TTTSAASSTLGGKSVSSNGRM 527
QY 456 RTEARGPEKTDQALARVISGNSRTL-----GDEVSVQVSALQVMQI 496

Db 528 ATSKTSSTSDLSKSVIFGNSSTVTPSPASISLTASPLSPVSDITSSEASSISSMLA 587
QY 497 TQSNPQANNEIR-----QKLTASVTKPPQFGYPYVQLSNDSTQKFIKLE 542
DB 588 SSSAPSDNNSTIASASLIVTKNSVSVSISSITSSSTTNSNLTATSTSLSNKATAR 647
QY 543 SLFAGSGRTAAEIKALSFEWNS 564
DB 648 SLSTSNATASAGNVPTGTFSSMS 669

RESULT 5
MSB2_YEAST
ID MSB2_YEAST STANDARD; PRT; 1306 AA.
AC P32334;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MSB2 protein (Multicopy suppressor of bud emergence 2).
GN MSB2 OR YGR014W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92383951; PubMed=1514328;
RA Bender A., Pringle J.R.;
RT "A Ser/Thr-rich multicopy suppressor of a cdc24 bud emergence
RT defect."
RL Yeast 8:315-323 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=97433481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII."
RL Yeast 13:1077-1090 (1997).
CC -!- PTM: O-glycosylated in the Ser/Thr-rich regions (Probable).
CC -!- SIMILARITY: SOME, TO YEAST HKR1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M77354; AAA34798.1; -.
DR EMBL; Z72799; CRA96997.1; -.
DR PIR; S25370; S25370.
DR GenOnline; 141326; -.
DR SGD; S0003246; MSB2.
DR GO; GO:0005887; C: integral to plasma membrane; IMP.
DR GO; GO:0005034; F: osmensor activity; IMP.
DR GO; GO:0000283; P: establishment of cell polarity (sensu Sacch. .); IGI.
DR GO; GO:0006970; P: response to osmotic stress; IMP.
KW Transmembrane; Glycoprotein; Repeat.
FT DOMAIN 698 816 7 X 17 AA TANDEM REPEATS.
FT REPEAT 698 714 1.
FT REPEAT 715 731 2.
FT REPEAT 732 748 3.
FT REPEAT 749 765 4.
FT REPEAT 766 782 5.
FT REPEAT 783 799 6.
FT REPEAT 800 816 7.
SQ SEQUENCE 1306 AA; 133114 MW; 67D5D984D5CA4A6D CRC64;

Query Match
Best Local Similarity 6.0%; Score 172; DB 1; Length 1306;
Matches 115; Conservative 127; Mismatches 277; Indels 163; Gaps 15;
```


Query Match 5.7%; Score 163.5; DB 1; Length 2090;
 Best Local Similarity 21.1%; Pred. No. 1.4;
 Matches 147; Conservative 88; Mismatches 262; Indels 199; Gaps 29;

QY	17	STSRASADVSTTATATPTP	-----PPTFDYKTAQTAVDYDTFTSTSLADIQAALVSLQD	71	
DB	940	SPMKQQLNFIAKRTPPVR	STAPASLASRAFLSORYEDLDVSTSSVSQSLDES-ED	998	
QY	72	AVTNIKDTAATDEE	-----	1054	
DB	999	ARTSCKDDEAVVQAPRHP	WVTPSPISQPSLLPHAAPFAKSHLVHGSSPGVMGTSA	93	
QY	94	TKNADAVKGAQITELAKY	ASDNOAILDSLKGLTSFDLLQALLOSVAANNKKAELL	150	
DB	1055	TSASKIIPGASDTMLATV	XHGAPSPHISAPQOLAAALRRQASQAPVNTLITES	1114	
QY	151	-----	KEMQDNPPVVGKTPAIAGLSLVDQTDAT-----	ATQTEKDGNAIRDA	191
DB	1115	TLKNVPQVVNQELKNPATP	--STANGSSVPPYSTAKTTPHPVLTPVAANAQKQGLNSL	1172	
QY	192	YFAGQN-AGGAVENAKSN	NSINIDSAKAIATAKTQIAEBAQKPPDPSPILQBAEQMVIQ	250	
DB	1173	KPSGPTFASQQLSSGDK	ASGTAKIETAVTSTPSASQGS--KPFSPSPSGTGFNGIIT	1229	
QY	251	AEKDLKNIKPADQ	-----SDVPNPGTIVGSGKQGGSSIGSRVSMLLDDAENETAS	--301	
DB	1230	PTFS-SNFTAAQATPST	KXESSQDFASFSSGGGSKPSYEALPESPSPSGITSASNTTPGEP	1288	
QY	302	-----	ILMSGFRMIHMPNTENPDSQAAQ-----QELAAQA-----	RAAKAAGDSDS	342
DB	1289	AASSRPPVAPSGTAL	STTSSEKLETPTPSKLGELLFPSLAGETLGSFGLRVGQA--DDST	1346	
QY	343	-----	AAALADAQKALEA-----ALGKAGQQQGLNALGATASAAVVSAGV	384	
DB	1347	KPTNKASSTLSLTQ	TKTSGVPSGNGFTAPPVLGHTEP-----	PVTSGATTISVA	1398
QY	385	PPAAASSIGSSV-KOLY	KTSK-STG--SDYKTOISAGYDAYKSINDAYGRARDAT----	436	
DB	1399	PPAATSTSTAVFGSL	PFVTSAGSSGVSIFSGTSLSAGKTSF-----SFGSQQTNSTVPPS	1453	
QY	437	-----	RDVINNVSTPALTRSVPRARTEARG-----PEKTOALARVI	473	
DB	1454	APPPTTAATPLTSP	FTFLSFGSLLSATTSPSPMSAGRSTEEATSSALPEK-----	1504	
QY	474	SGNSRTLGDVYSOVS	ALQSMVQITQSNPOANNEIRQKLTSVATKPPQFGYPVYQLSND	533	
DB	1505	PGDSEVSASASLLE	EQQSA-QLPQAPPQTSQ-----SVKKEFVLQAPAV--SNSG	1552	
QY	534	TQKFIAKLSLPAEG	-----SRTAAEIKALSF	560	
DB	1553	TAASTSLVALSAAE	ATPATTGVDFARTEAVPPASSF	1588	

RESULT 12
 MINT_MOUSE
 ID MINT_MOUSE STANDARD; PRT; 3644 AA.
 AC G62504; Q80TN9; Q99PS4; Q9QZW2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Mx2-interacting protein (SMART/HDAC1 associated repressor protein).
 GN MINT OR SHARP OR KIA00929.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN
 RP
 RC
 RP
 RX
 RA

SEQUENCE OF 49-3644 FROM N.A. (ISOFORM 1), FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, DNA-BINDING, AND INTERACTION WITH MSX2. TISSUE=Testis;
 MEDLINE=99379811; PubMed=10451362;
 Newberry E.P., Latifi T., Towler D.A.;

RT "The RRM domain of MINT, a novel msx2 binding protein, recognizes and
 RT regulates the rat osteocalcin promoter.";
 RL Biochemistry 38:10678-10690(1999).
 RN [2]
 RP SEQUENCE OF 1-112 FROM N.A.
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P., Fukuda S.,
 RA Hashizume W., Hayashida K., Hirozane T., Hori F., Inotani K.,
 RA Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y., Kondo S., Konno H.,
 RA Koya S., Miyazaki A., Murata M., Nakamura M., Nomura K., Numazaki R.,
 RA Ohno M., Ohsato N., Saito R., Sakazume N., Sano H., Sasaki D.,
 RA Sato K., Shibata K., Shiraki T., Tagami M., Takeda Y., Waki K.,
 RA Watahiki A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 69-3644 FROM N.A. (ISOFORM 2), AND VARIANTS THR-348;
 RP PHE-762; PHE-773 AND LEU-933.
 RC STRAIN=ICR; TISSUE=Brain;
 RA Sakamoto T., Gotou T., Isagawa Y., Mimura H., Kimura K., Kawaichi M.;
 RT "MINT/spen negatively regulates Notch signaling by inhibiting RBP-
 RT J/Su(H) activity.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 318-578 FROM N.A.
 RC TISSUE=Cochlea;
 RX MEDLINE=97237053; PubMed=9119401;
 RA Crozet F., El-Amraoui A., Blanchard S., Lenoir M., Ripoll C., Vago P.,
 RA Hamel C., Fizaes C., Levi-Acobas F., Depetris D., Mattei M.-G.,
 RA Weil D., Pujol R., Petit C.;
 RT "Cloning of the genes encoding two murine and human cochlear
 RT unconventional type I myosins.";
 RL Genomics 40:332-341(1997).
 RN [5]
 RP SEQUENCE OF 2598-3644 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22579291; PubMed=12693553;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
 RA Nakajima D., Nagase T., Ohara O., Koga H.;
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
 RT cDNAs identified by screening of terminal sequences of cDNA clones
 RT randomly sampled from size-fractionated libraries.";
 RL DNA Res. 10:35-48(2003).
 RN [6]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=22261914; PubMed=12374742;
 RA Oswald P., Kostecka U., Astrahantseff K., Bourteelle S., Dillinger K.,
 RA Zechner U., Ludwig L., Wilda M., Hameister H., Knoechel W., Liptay S.,
 RA Schmid R.M.;
 RT "SHARP is a novel component of the Notch/RBP-Jkappa signalling
 RT pathway.";
 RL EMBO J. 21:5417-5426(2002).
 RN [7]
 RP FUNCTION, AND TISSUE SPECIFICITY.
 RX MEDLINE=22483652; PubMed=12594956;
 RA Kuroda K., Han H., Tani S., Tanigaki K., Tun T., Furukawa T.,
 RA Taniguchi Y., Kurooka H., Hamada Y., Toyokuni S., Honjo T.;
 RT "Regulation of marginal zone B cell development by MINT, a suppressor
 RT of Notch/RBP-J signaling pathway.";
 RL Immunity 18:301-312(2003).
 CC -!- FUNCTION: Essential corepressor protein, which probably regulates
 CC different key pathways such as the Notch pathway. Negative
 CC regulator of the Notch pathway via its interaction with RBPSUH,
 CC which prevents the association between NOTCH1 and RBPSUH, and
 CC therefore suppresses the transactivation activity of Notch
 CC signaling. Blocks the differentiation of precursor B cells into
 CC marginal zone B cells. Probably represses transcription via the
 CC recruitment of large complexes containing histone deacetylase
 CC proteins. May bind both to DNA and RNA.
 CC -!- SUBUNIT: Interacts with NCOR2, HDAC1, HDAC2, RBBP4, MBD3 and
 CC TALL1. Interacts with the nuclear receptors RAR and PPARG.
 CC Interacts with RAR in absence of ligand. Bind to the steroid
 CC receptor RNA coactivator SRA (By similarity). Interacts with MSX2.
 CC Interacts with RBPSUH; this interaction may prevent the

CC interaction between RBPSUH and NOTCH1.
 CC -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q62504-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q62504-2; Sequence=VSP_008564;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Highly expressed in testis. Expressed at lower
 CC level in brain, lung, spleen, liver and kidney. Weakly expressed
 CC in cardiac and skeletal muscles and ovary. In spleen, it is
 CC expressed in follicular B-cells, while it is weakly expressed in
 CC marginal zone B-cells.
 CC -!- DOMAIN: The RID domain mediates the interaction with nuclear
 CC receptors.
 CC -!- DOMAIN: The SPOC domain, which mediates the interaction with
 CC NCOR2, is essential for the repressive activity (By similarity).
 CC -!- SIMILARITY: Belongs to the Spen family.
 CC -!- SIMILARITY: Contains 1 RID (receptor interacting) domain.
 CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
 CC -!- SIMILARITY: Contains 1 SPOC domain.
 CC -!- CAUTION: Ref.4 sequence differs from that shown due to multiple
 CC frame shifts and conflicts that create stop codons.
 CC -!- CAUTION: Ref.5 sequence differs from that shown due to what seems
 CC to be the presence of intronic sequence in the cDNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; BX726481; -- NOT ANNOTATED CDS.
 DR EMBL; AF156529; AAD55931.1; ALT_INIT.
 DR EMBL; AF156529; BAB32786.1; --
 DR EMBL; Z78160; CAB01562.1; ALT_SEQ.
 DR EMBL; AK122402; BAC65684.2; ALT_SEQ.
 DR MGD; MGI:1391706; Mint.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; Rtm; 3.
 DR SMART; SM00360; RRM; 3.
 DR PROSITE; PS01002; RRM; 4.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR PROSITE; PS00917; SPOC; 1-1;
 KW Transcription regulation; Repressor; Nuclear protein; DNA-binding;
 KW RNA-binding; Repeat; Coiled coil; Alternative splicing; Polymorphism.
 FT DOMAIN 1 574
 FT DOMAIN 6 81
 FT DOMAIN 336 416
 FT DOMAIN 439 514
 FT DOMAIN 518 590
 FT DOMAIN 559 575
 FT DOMAIN 822 850
 FT DOMAIN 1185 1206
 FT DOMAIN 1509 1544
 FT DOMAIN 1607 1627
 FT DOMAIN 2216 2704
 FT DOMAIN 3478 3644
 FT DOMAIN 2138 2462
 FT DOMAIN 2706 2845
 FT DOMAIN 125 277
 FT DOMAIN 236 326
 FT DOMAIN 648 721
 FT DOMAIN 702 832
 FT DOMAIN 2101 2233
 FT DOMAIN 2377 2518
 FT DOMAIN 2950 3475
 FT VARSPLIC 618 640
 FT VARIANT 348 348
 FT I -> T.
 FT /FTid=VSP_008564.

```
FT VARIANT 762 762 S -> F.
FT VARIANT 773 773 S -> F.
FT VARIANT 933 933 R -> L.
FT VARIANT 754 754 R -> G (IN REF. 3).
FT CONFLICT 1524 1524 D -> A (IN REF. 3).
FT CONFLICT 1560 1560 H -> Y (IN REF. 3).
FT CONFLICT 1570 1570 F -> L (IN REF. 3).
FT CONFLICT 1574 1574 Q -> G (IN REF. 3).
FT CONFLICT 1609 1609 R -> R (IN REF. 3).
FT CONFLICT 1659 1659 I -> V (IN REF. 3).
FT CONFLICT 1669 1669 S -> F (IN REF. 3).
FT CONFLICT 1705 1705 V -> A (IN REF. 3).
FT CONFLICT 1815 1815 A -> V (IN REF. 3).
FT CONFLICT 2097 2097 G -> A (IN REF. 3).
FT CONFLICT 2201 2202 MISSING (IN REF. 3).
FT CONFLICT 2322 2322 A -> V (IN REF. 3).
FT CONFLICT 2385 2385 P -> Q (IN REF. 3).
FT CONFLICT 2502 2502 R -> K (IN REF. 3).
FT CONFLICT 2505 2505 E -> K (IN REF. 3).
FT CONFLICT 2519 2519 D -> N (IN REF. 3).
FT CONFLICT 2554 2554 T -> S (IN REF. 3).
FT CONFLICT 2679 2679 LVSTDPAGPVN -> VGEHPWARD (IN REF. 3).
FT CONFLICT 3010 3010 L -> P (IN REF. 3 AND 5).
SQ SEQUENCE 3644 AA; 398750 MW; 9C7EC49A81A7DA4A CRC64;

Query Match 5.7%; Score 163; DB 1; Length 3644;
Best Local Similarity 18.1%; Pred. No. 2.9;
Matches 115; Conservative 105; Mismatches 274; Indels 140; Gaps 16;

Qy 20 RSADVSTATATPPPTFDYKTOAQATYDTFTSTSLADIQAALVSLQDVTNKT 79
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1620 KEABQKEPETHKTPPEAAETKEPEK-----APVSAGLPAVITVVTPEPA 1667
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 80 AATDETAIAAEWETKNDAVKVGAQITELAKYASDNOAI-----LDSLGL 126
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1668 SSAPEKAEEAEAPSPAGEKPAEPVSEETKLSEPVSVFVEQPRQSDVPFGEDSRDSQ 1727
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 127 TSFDLLQALQSVANNKAAELKEMQDNVPVPGKTPALAQSLVD----- 172
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1728 DS-----AALAPSAPOESAATDAVPCVNAEPLTGTGTVSQVESVDPKPSPLSKLTQ 1782
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 173 -QTDATATQIEK-----DGNALRDYFAQNASGAVENAKSNNSINIDSAKAATA 224
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1783 RSEABEGKVKPDTTPTSTEDATQNAVGAEOPPASEDVEANPPVAAKDRKTNKSKRS 1842
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 225 KTOIAEAQKPPDSPILQEAEE-----QMVIOAEKDLKNIKPADG 263
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1843 KTSVQAAASVVEKPVTRKSERIDREKLKRSSSPRGEAQKLELQWEAEKITRASKSSG 1902
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 264 SDVNPPTGVGSKQOGSSIGSRVSMLLDDEANE-----TASILMSGFRQMTIM 313
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1903 GDTHEPEPESPLSRSRNRNVSATVMTDHSRSPAKEPVEQPRVTRKRLREQLQEAATVP 1962
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 314 FNT-----ENPDQAAQOELAAQARAACA-----AGDDSAALADAKALEAA 357
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1963 PTTFRGPPKTRRAEDGHEKKEPATETPRPAEGWSPRSQSKAAAA-----GP 2013
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 358 LGKAGQOQGIINALQIAAAVWSAGVPPA-----AASSIGSVKLYKTSKSGDYKTQ 413
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2014 QGKGRNEQKVEAAEAQAQASTREGNPKSRGEREAASEPKRDRDPSTDSKSPDTPFVE 2073
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 414 ISAGVDAYSNDAYGRARNDAITRVINNVSTPALTRSVPRPAREGPEKTDQALAVI 473
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2074 VLKRPPEKTYKSKRGAR--STRSGMDRAHQRSLERMAAAGQAADKE----- 2121
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 474 SGNRTLDGVYSQVSAQSVMTQSNPQ-ANN-----EIRQKLTSAVTKPPQFGYFV 527
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2122 -----AGPAASPOESSPQKSGSSPOLANNPADPDRABEESASASTAPPETQLAR 2175
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 528 QLSNDSTOKFTAKL-----ESLFAEGSRITAAEIK 557
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2176 QIELEQAVQNIAKLPEPSAAAASKGTATATATAA 2209
: : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 13

```
TALA_DICDI
ID TALA_DICDI STANDARD; PRT; 2492 AA.
AC P54633;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Filopodin (Talin homolog).
GN TALA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RE MEDLINE=95213284; PubMed=7698984;
RA Kreitmeier M., Gerisch G., Heizer C., Mueller-Taubenberger A.;
RT "A talin homologue of Dictyostelium rapidly assembles at the leading
edge of cells in response to chemoattractant.";
RL J. Cell Biol. 129:179-188(1995).
RN [2]
RP REVISIONS TO 2076; 2080; 2186 AND 2230.
RA Mueller-Taubenberger A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Actin-binding protein that may be involved in the
control of cell motility and chemotaxis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; rapidly assembles at the
leading edge of cells in response to chemoattractant.
CC -!- SIMILARITY: Contains 1 FERM domain.
CC -!- SIMILARITY: Contains 1 I/LWEQ domain.
CC PARTIAL, TO YEAST ROD1.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; U14576; AAC4586.2; -.
DR DictyBase; DB0001399; talA.
DR InterPro; IPR002099; Band 4.1.
DR InterPro; IPR002558; ILWEQ.
DR Pfam; PF00373; Band 4.1; 1.
DR Pfam; PF01608; ILWEQ; 1.
DR ProDom; PD011820; ILWEQ; 1.
DR SMART; SM00295; B4.1; 1.
DR SMART; SM00307; ILWEQ; 1.
DR PROSITE; PS00660; FERM_1; 1.
DR PROSITE; PS00661; FERM_2; 1.
DR PROSITE; PS00057; FERM_3; 1.
DR PROSITE; PS00945; ILWEQ; 1.
KW Structural protein; Cytoskeleton; Actin-binding.
FT DOMAIN 84 365
FT DOMAIN 2298 2492
FT SEQUENCE 2492 AA; 268850 MW; 4B82F6BE6DE7E752 CRC64;
SQ SEQUENCE 2492 AA; 268850 MW; 4B82F6BE6DE7E752 CRC64;

Query Match 5.7%; Score 161.5; DB 1; Length 2492;
Best Local Similarity 20.6%; Pred. No. 2.1;
Matches 121; Conservative 99; Mismatches 263; Indels 103; Gaps 20;

Qy 6 KLGIASSNSSSSRSADVSTTATPTPPPTFDYKTOAQATYDTFTSTSLADIQA 65
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1982 KLLIVSSQLVGSRCFQEHLSATTKV-----ANTVSSLVKTAKIASVLAD 2029
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 66 LVSLQDVTNKTATDEBETAIAEWETKNDAVKGAQITELAKYASDNOAILDSLGL 125
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2030 TTSQQDILSASKALSISSQQMWLA---TKDAQRFKDA--TAFRSLGSAEVAEAVGQ 2083
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 126 -LTSFDLLQALQSVANNKAAELKEMQDNVPVPGKTPALAQSLVDQTDAT----- 177
: : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 2084 FLTS-----VYTAISDRGK-----IKELEKIVQVANHVKPDVLSNKNKDATAEIFAQS 2133
QY 178 ATQIEKDGNAIRDYFAGQNASGAVENAKSNNSINISDAKAAIATAKTQIAEAKKFPD 237
Db 2134 ARDLAKSSIEIIVTSYSSQSLVKSSQAVVNVQSFINSKGVIALLGNGNDLKSCK-- 2190
QY 238 SPILQEAQWVCAEKDLKNIKPAQSDVFNPGTIVGGSKQGGSSIGSIRVSMLLDDARN 297
Db 2191 --VLENVQTQTDMLALQCVKQDQ-----KNGSTSIADATRSID-----RVHSVV----- 2235
QY 298 ETASILMSGFROMIHMFENTENPDSQA-AQELAAQARAARAAAGDDSAALADAKALEA 356
Db 2236 --TLKSLPGQNVVEEDNVLEALAEDELASACR-----SIEETAKLIA 2282
QY 357 ALGKAGQQGILNALQIASAAVVSAGVPPAAASSIGSVKOLYKTSK---STGSDYKQT 413
Db 2283 ARPQSKSKNGKLDAGVAAATIVDASSAIKAVAKLVNSAAVQSKRRERQIASGSGVYKAD 2342
QY 414 -----ISA-----GYDAYK-----SINDAYGRANDATEDVINNV-STPALTRSVPRART 457
Db 2343 PTWSNGLISAAKGVGAATHRLVEAAKMSATGKAEEELIATARSVAATAALIVSARAKS 2402
QY 458 EARGPEKTDQALARVISGNSRILGDIYVSOVSALQSVMLQTNPOANNEEIRKLTSAVT 517
Db 2403 ---GDYQSOAHSLSHTAARQVASATSLVAAKAAATIFDQOQEEEGE----- 2449
QY 518 KPPQGYPVVQVSNDSQTKFIAKLSLFAEGSRTAAEIKALSFETN 563
Db 2450 ---QNFYTSKVKLEQEQMKILKELETRARQMLNSRKQNYKN 2492

RESULT 14
DAN4 YEAST
ID DAN4 YEAST STANDARD; PRT; 1161 AA.
AC P47179;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell wall protein DAN4 precursor.
GN DAN4 OR YUK151C OR J2223.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Scarce T.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RX REGULATION.
RA MEDLINE=21113168; PubMed=11160904;
RA Cohen B.D., Sertif O., Abramova N.B., Davies K.J., Lowry C.V.;
RT "Induction and repression of DAN1 and the family of anaerobic
RT mannoprotein genes in Saccharomyces cerevisiae occurs through a
RT complex array of regulatory sites.";
RL Nucleic Acids Res. 29:799-808(2001).
CC -!- FUNCTION: Component of the cell wall (By similarity).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -!- PTM: Extensively O-glycosylated (Potential).
CC -!- SIMILARITY: Belongs to the SRP1 / TIPI family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z49651; CAA89684.1; -.
CC FIC; S57180; S57180.
CC GermOnline; 141984; -.
DR
```

```
DR SGD; S0003912; DAN4.
DR GO; GO:0009277; C:cell wall (sensu Fungi); IDA.
DR InterPro; IPR000992; SRP1_TIP1.
DR Pfam; PF00660; SRP1_TIP1; 1.
DR PROSITE; PS00724; SRP1_TIP1; 1.
KW Cell wall; Glycoprotein; Membrane; GPI-anchor; Signal; Lipoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 1146 CELL WALL PROTEIN DAN4.
FT PROPEP 1147 1161 REMOVED IN MATURE FORM (POTENTIAL).
FT LIPID 1146 1146 GPI-anchor amidated asparagine
FT (Potential).
SQ SEQUENCE 1161 AA; 118358 MW; 7954C15D69F0CA58 CRC64;

Query Match 5.6%; Score 160; DB 1; Length 1161;
Best Local Similarity 20.1%; Pred. No. 0.98;
Matches 118; Conservative 92; Mismatches 274; Indels 104; Gaps 21;

QY 11 SSSSSSSRSADVDSTTATPTPPPTFDDYKTCQAQYDTIFTSTSLADIQAALVSL- 69
Db 252 STSTPTPTTPTTPTTSTTAPTSTTSTTSTTSTTSTTSTTSTTSTTSTTSTTSSASSVI 311
QY 70 -----QDAVINIKDIAATDEETAIAAEWETKNADAVKGAQITELAKYADNQAI 119
Db 312 STTATSTTFLTPATSTASTDHTTSSVS---TTFATTSATTTT-----SDTYS 362
QY 120 LDSLGKLTFLDLQALLQSVANNKAAELKEMQDNVVPKGTALIAQSLVDQDAT-A 178
Db 363 SSSPQVTS-----SAEPTTVSEVTSSVEPTRSQ---VTSSAEPTTVSEPTSSVEPTRS 414
QY 179 TCIEKGNAIRDYFAGQNASGAVENAKSNNSINISDAKAAIATAKTQIAEAKKFPDS 238
Db 415 SQVTSSAEPTTVSEF-----TSSVEPTRSQVTS-----SAEPTTVSEPTSSVEPTRSS 463
QY 239 PILOEAEQWVCAEKDLKNIKPAQSDVFNPG---TTVGSKQGGSSIGSIRVSMLLDDAE 296
Db 464 QVTSSAEPTTV--SEFTSSVEPTRSSQVTSAAEPTV--SEFTSSVEPTRSSQVTSAAE 518
QY 297 NETASILMSGFROM-IHMFENTENPDSQ---AAQELAAQARAARAAAGDDSAALADAKA 353
Db 519 PTTVSEVTSSVEPTRSSQVTTTEPVSPFGSTFEITSSAEPLSFSAKATTSAESISSNQIT 578
QY 354 LEALGKAGQQGILNALQIASAAVVSAGVPPAAASSIGSVKOLYKTSKSGDYKQTQ 413
Db 579 ISSL-----IVSSVITS-----SSEIPSSIEVL--TSSGSSSVSEPT 614
QY 414 ISAGYDAYKSIDAYGRANDATEDVINNVSTPA--LTRSVPRARTEARGPEKTDQALAR 471
Db 615 SLVGPSDESISSTESLSATSTFTTSVVSSSKAADFFTRSTVSAKSD----- 661
QY 472 VISGNSRSLGDIYVSOVSALQSVMLQTNPOANNEEIRKLTSAVKPPQGYPVVQVLSN 531
Db 662 -VSGNSSTQSTTFF-----ATPSTPLAVSSTVVTSTDSVS--PNTPFSEISSP 708
QY 532 DSTQKFIKLSLFAE--GSRTAAEIKALSPETNSLFIQQVLVNIQSL 577
Db 709 ESSTAITSTSTSFIAERTSSLYLSSNNMSSFTLSTFTVTSQIVSVSFSM 756

RESULT 15
SPAP STRMU
ID SPAP STRMU STANDARD; PRT; 1562 AA.
AC P23504;
DT 01-NOV-1991 (Rel. 20, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell surface antigen I/II precursor.
GN SPAP OR SMU.610.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 23, 2004, 18:47:55 ; Search time 43.9152 Seconds
(without alignments)
4188.690 Million cell updates/sec

Title: US-10-608-559-4
Perfect score: 2850
Sequence: 1 MSLADKLGIASSNSSSTR.....SLFIOQLVWIGSLVGYLQ 583

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum Match 100%
Maximum DB seq length: 2000000000
Listing first 45 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviro.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2844	99.8	651	16 Q9Z7H7	Q9Z7H7 chlamydia p
2	1870	65.6	715	2 Q46166	Q46166 chlamydia p
3	1129	39.6	636	16 Q824X3	Q824X3 chlamydia p
4	843	29.6	647	16 Q84627	Q84627 chlamydia t
5	224.5	7.9	2310	16 Q8CMU7	Q8CMU7 staphylococ
6	210	7.4	2055	2 Q85472	Q85472 abiotrophia
7	206	7.2	6713	16 Q99U54	Q99U54 staphylococ
8	205	7.2	1822	2 Q07290	Q07290 streptococ
9	205	7.2	2016	5 Q9BIT0	Q9BIT0 plectreureys
10	202	7.1	6713	16 Q931R6	Q931R6 staphylococ
11	200.5	7.0	9439	16 Q8CP76	Q8CP76 staphylococ
12	200	7.0	1313	5 Q9NE28	Q9NE28 leishmania
13	200	7.0	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
14	199.5	7.0	2275	16 Q8NUJ3	Q8NUJ3 staphylococ
15	199	7.0	973	16 Q8XDQ4	Q8XDQ4 escherichia
16	198	6.9	762	16 Q89IU6	Q89IU6 bradyrhizob

17	197	6.9	2478	2 Q9RL69	Q9RL69 staphylococ
18	197	6.9	2478	2 Q9LCH2	Q9LCH2 staphylococ
19	195.5	6.9	2271	16 Q99QY4	Q99QY4 staphylococ
20	194	6.8	2283	2 Q8VQ99	Q8VQ99 staphylococ
21	192	6.7	2167	2 Q845L8	Q845L8 mycoplasma
22	191.5	6.7	806	16 Q88Y67	Q88Y67 lactobacill
23	190.5	6.7	1545	16 Q9RDQ1	Q9RDQ1 streptomyce
24	190	6.7	856	5 Q86RM6	Q86RM6 plasmodium
25	189	6.6	855	5 Q86DH2	Q86DH2 plasmodium
26	189	6.6	1063	16 Q9CH86	Q9CH86 lactococcus
27	189	6.6	1365	2 Q49525	Q49525 mycoplasma
28	188	6.6	803	5 Q86RN4	Q86RN4 plasmodium
29	187.5	6.6	1795	16 Q9LCU9	Q9LCU9 staphylococ
30	187.5	6.6	2481	16 Q99QR6	Q99QR6 staphylococ
31	187.5	6.6	4776	16 Q97P71	Q97P71 streptococ
32	186	6.5	3360	16 Q88XB6	Q88XB6 lactobacill
33	185.5	6.5	1336	16 Q9ACV2	Q9ACV2 streptomyce
34	184.5	6.5	642	2 Q7X4Y6	Q7X4Y6 rhodospiril
35	184.5	6.5	855	5 Q86RM1	Q86RM1 plasmodium
36	183	6.4	625	5 Q86RN6	Q86RN6 plasmodium
37	182.5	6.4	3072	2 Q939N5	Q939N5 streptococ
38	182	6.4	842	5 Q86RN1	Q86RN1 plasmodium
39	182	6.4	3381	2 Q9KX33	Q9KX33 streptococ
40	181.5	6.4	2541	13 Q8AM10	Q8AM10 gallus gall
41	180.5	6.3	1327	2 Q9X7M2	Q9X7M2 staphylococ
42	180	6.3	2178	2 Q9KWR3	Q9KWR3 streptococ
43	180	6.3	2541	4 Q86YD0	Q86YD0 homo sapien
44	180	6.3	2550	4 Q9UPX3	Q9UPX3 homo sapien
45	179.5	6.3	641	16 Q9KS57	Q9KS57 vibrio chol

ALIGNMENTS

RESULT 1

ID	Q9Z7H7	PRELIMINARY;	PRT;	651 AA.
AC	Q9Z7H7			
DT	01-MAY-1999 (TREMELrel. 10, Created)			
DT	01-MAY-1999 (TREMELrel. 10, Last sequence update)			
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)			
DE	CHLPN 76 kDa HOMOLOG 1 (CT622) (Hypothetical protein).			
GN	CPN0728 OR CPJ0728 OR CP0018 OR CPB0756.			
OS	Chlamydia pneumoniae (Chlamydia pneumoniae).			
OC	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.			
OX	NCBI_TaxID=833558;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CNL029;			
RX	MEDLINE=9920606; PubMed=10192388;			
RA	Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,			
RA	Olinger L., Grimwood J., Davis R.W., Stephens R.S.;			
RT	"Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.";			
RL	Nat. Genet. 21:385-389(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ARJ9;			
RX	MEDLINE=20150255; PubMed=10684935;			
RA	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,			
RA	White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,			
RA	Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,			
RA	Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,			
RA	Eisen J., Fraser C.M.;			
RT	"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia			
RL	pneumoniae AQ39.";			
RT	Nucleic Acids Res. 28:1397-1406(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=J138;			
RX	MEDLINE=20330349; PubMed=10871362;			
RA	Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,			
RA	Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;			
RT	"Comparison of whole genome sequences of Chlamydia pneumoniae J138			

```

RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TW-183;
RA Geng M.M., Schuhmacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE001654; AAL18867.1; -
DR EMBL; AE002165; AAF37914.1; -
DR EMBL; AP002547; BAA98935.1; -
DR EMBL; AE017159; BAA98685.1; -
DR PIR; D72042; D72042.
DR PIR; E86581; E86581.
DR PHCI-2DPAGE; Q927H7; -.
DR TIGR; CP0018; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 651 AA; 68217 MW; 47AE6C3F2FF0123 CRC64;

Query Match 99.8%; Score 2844; DB 16; Length 651;
Best Local Similarity 99.8%; Pred. No. 2.3e-124;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 60
Db 69 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 128
QY 61 DIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 120
Db 129 DIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 188
QY 121 DSLGKLTSEFLLQALQLQSVANNKKAELLKEMQDNPVPGKTPAIAQSLVDQTDATATQ 180
Db 189 DSLGKLTSEFLLQALQLQSVANNKKAELLKEMQDNPVPGKTPAIAQSLVDQTDATATQ 248
QY 181 IEKGNAIRDAYFAGQNASGAVENAKSNNSISNDSAKAAIATAKTQIAEAQKFPDPSPI 240
Db 249 IEKGNAIRDAYFAGQNASGAVENAKSNNSISNDSAKAAIATAKTQIAEAQKFPDPSPI 308
QY 241 LQAEQWVIOAEKDLKNIKPADGSDVPNGTGVGSKQGGSSIGSVMLDDAENETA 300
Db 309 LQAEQWVIOAEKDLKNIKPADGSDVPNGTGVGSKQGGSSIGSVMLDDAENETA 368
QY 301 SILMSGFRQMIHMENTENPDSQAQQLAQAARAAKAGDSDSAAALADAKALEAALGK 360
Db 369 SILMSGFRQMIHMENTENPDSQAQQLAQAARAAKAGDSDSAAALADAKALEAALGK 428
QY 361 AGQQGGILNALGQIASAAVVSAGVPPAAASIGSSVQLYKTSKSGDYKTIQISAGYDA 420
Db 429 AGQQGGILNALGQIASAAVVSAGVPPAAASIGSSVQLYKTSKSGDYKTIQISAGYDA 488
QY 421 YKSNIDAYGRARNDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVIGNSRTL 480
Db 489 YKSNIDAYGRARNDATRDVINNVSTPALTRSPVPRARTEARPEKTDQALARVIGNSRTL 548
QY 481 GDVYSQVSALQSVQMIIQSNPQANNEIRKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 540
Db 549 GDVYSQVSALQSVQMIIQSNPQANNEIRKLTSAVTKPPQFGYPYVQLSNDSTQKFIK 608
QY 541 LESLFAEGSRTAARAEIKALSFNTSLFIQOVLNIGLSYGLQ 583
Db 609 LESLFAEGSRTAARAEIKALSFNTSLFIQOVLNIGLSYGLQ 651

RESULT 2
Q46166 ID Q46166 PRELIMINARY; PRT; 715 AA.
AC Q46166;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)

76 kDa protein.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94156481; PubMed=7509320;
RA Perez-Melgosa M., Kuo C.-C., Campbell L.;
RT "Isolation and characterization of a gene encoding a Chlamydia
RT pneumoniae 76-kilodalton protein containing a species-specific
RT epitope.";
RL Infect. Immun. 62:880-886(1994).
DR EMBL; L23921; AAA23117.1; -
DR PIR; I40729; I40729.
SQ SEQUENCE 715 AA; 76626 MW; 3F01C9A59DE964C5 CRC64;

Query Match 65.6%; Score 1870; DB 2; Length 715;
Best Local Similarity 99.7%; Pred. No. 4.4e-79;
Matches 385; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 60
Db 325 MSLADKLGIASSNSSSTSRSDVSTTATATPTPPPTFDYKTAQAYDTIFTSTSLA 384
QY 61 DIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 120
Db 385 DIQAALVSLQDAVTNICKDTAATDEETAIAAEWETKNADAVKGAQITELAKYASDNQAIL 444
QY 121 DSLGKLTSEFLLQALQLQSVANNKKAELLKEMQDNPVPGKTPAIAQSLVDQTDATATQ 180
Db 445 DSLGKLTSEFLLQALQLQSVANNKKAELLKEMQDNPVPGKTPAIAQSLVDQTDATATQ 504
QY 181 IEKGNAIRDAYFAGQNASGAVENAKSNNSISNDSAKAAIATAKTQIAEAQKFPDPSPI 240
Db 505 IEKGNAIRDAYFAGQNASGAVENAKSNNSISNDSAKAAIATAKTQIAEAQKFPDPSPI 564
QY 241 LQAEQWVIOAEKDLKNIKPADGSDVPNGTGVGSKQGGSSIGSVMLDDAENETA 300
Db 565 LQAEQWVIOAEKDLKNIKPADGSDVPNGTGVGSKQGGSSIGSVMLDDAENETA 624
QY 301 SILMSGFRQMIHMENTENPDSQAQQLAQAARAAKAGDSDSAAALADAKALEAALGK 360
Db 625 SILMSGFRQMIHMENTENPDSQAQQLAQAARAAKAGDSDSAAALADAKALEAALGK 684
QY 361 AGQQGGILNALGQIASAAVVSAGVPP 386
Db 685 AGQQGGILNALGQIASAAVVSAGVLP 710

RESULT 3
Q824X3 ID Q824X3 PRELIMINARY; PRT; 636 AA.
AC Q824X3;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN CCA00015.
OS Chlamydia pneumoniae.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RA MEDLINE=22569155; PubMed=12682364;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J., Holtzapple E., Khouri H., Federova M.B., Carty H.A.,
RA Umayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavoil P.M.,
RA Fraser C.M.;
RT "Genome sequence of Chlamydia pneumoniae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae.";
```


KW	Complete proteome.	DR	T31110; T31110.
SQ	SEQUENCE 2310 AA; 232314 MW; AC96B49EAB754B88 CRC64;	DR	GO:0016020; C:membrane; IEA.
	Query Match 7.9%; Score 224.5; DB 16; Length 2310;	DR	InterPro; IPR005877; Gpos_YsIRK.
	Best Local Similarity 17.2%; Pred. No. 0.051;	DR	Pfam; PF04650; YsIRK_signal; 1.
	Matches 110; Conservative 149; Mismatches 273; Indels 109; Gaps 16;	DR	TIGRFAMs; TIGR01168; YsIRK_signal; 1.
		FT	NON TER 2055 2055
		SQ	SEQUENCE 2055 AA; 215642 MW; 9699C11DD9E93E2FD CRC64;
QY	2 SLADKLGIASNSSGTSRSADVSTTATATPTPPPTFDYKTAQATAYTIITFTSLAD 61		
Db	1025 SLSESTSTSDSASTSESDNS-TSTSLSESTSTSLSESTSTSTSDSASTSESD 1082		
QY	62 IQAALVSLQDAV-TNKKD---TAATDETAIAAEWETKNADAVKVGACITELAKYA--- 113		
Db	1083 SNSASTSLRESTSTSLSDSTSTSDSASTSTSESDNSASTSLSGSSTSVSDSTAST 1142		
QY	114 -----SDNQAILDSGLKTSFDLLQ-----AALLOSANNKKAELLKEMQ 154		
Db	1143 SESASTSTSDNSSTSTSLSESTSTSLSESTSTSTSDSASTSMVSNSASTSLSEST 1202		
QY	155 DNPVVGKTPALAAQL-----VQDTATATQIEKGNNAIRDAYFAGNAGSAGVENAKSN 208		
Db	1203 STSVSDSTSTSEASASTSESDNSASTSLSESTSTSVSDS-----TSTSTSDSASTS 1257		
QY	209 NSISNIDSAAKAI-----ATAKTAIAEAAQKFPDPSPILOEAEQMWI 249		
Db	1258 TVSDNSASTSLSESTSTSDSTSTSDSASTSTSVSES-----SSTSTSIERSLS 1311		
QY	250 QAEKLNKIPADGSDVNPPTVGGSKQGSIGSIRVSMLLDDA----- 295		
Db	1312 TSDSDSKMSTSESA---STSVSDSEASTSESTSTSLSDSTSTSTSDSASTSE 1368		
QY	296 --ENETASTILMGFGFMHMFENTENPDSQAQELAAQARAAGAAGDSDAAALDAQKA 353		
Db	1369 SDNSNSTSKMSLSLSTSVSDSTSTSDSASTSTSVSDNSASTSLSDSTSTSDSTA 1428		
QY	354 L-----EALGKAGQQGILNALGQIASAAVVSAGVPPAAASSIGSSVKLYKTSK 404		
Db	1429 STSTSVSDNSASTSLSESTSTSLSDSTSVSTSDSTSTSESDNSSTSTSLSESTSL 1488		
QY	405 STGSDYKTKISAGYDAYKINDAYGRANDATRDVINNVSTPALTRSVPRARTARGPEK 464		
Db	1489 SUSTSTSDSASTSASESDNSASTSLSDSTSTSLSDSTSTSESTSTSESTSESES-AS 1547		
QY	465 TPQALARVIGNSRSLGVDYVSVALQSVMQIT-QSNPQANNEIRQKLTSAVTKPPQFG 523		
Db	1548 TSKKLSASASTSEASASTSTSTSDSASTSTSESDNSASTSMSESLSTSV----- 1600		
QY	524 YPVVQLSNSTQKFIKLSFAEGSRTAAETIKALSPETNS 564		
Db	1601 -----SDSTSMSTSEASTSTSESESNSESTSLSDSTST 1634		
RESULT 6			
O85472	PRELIMINARY; PRT; 2055 AA.		
ID	O85472		
AC	O85472; AC		
DT	01-NOV-1998 (TrEMBLrel. 08, Created)		
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
GN	Extracellular matrix binding protein (Fragment).		
DE	EMB.		
OS	Abiotrophia defectiva.		
OC	Bacteria; Firmicutes; Lactobacillales; Aerococcaceae; Abiotrophia.		
OX	NCBI_TaxID=46125;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=VNS-47;		
RE	MEDLINE=99081722; PubMed=9864195;		
RA	Manganelli R., Van de Rijs I.;		
RT	"Characterization of emb, a gene encoding the major adhesin of		
RT	Streptococcus defectivus."		
DR	Infect. Immun. 67:50-56 (1999).		
DR	EMBL; AF067776; AAD03320.1; .		

[illegible]

DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Epf* protein.
GN	EPF*
OS	Streptococcus suis.
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC	Streptococcus.
OX	NCBI_TaxID=1307;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=1890;
EX	MEDLINE=33328288; PubMed=8335363;
RT	Smith H.E., Reek F.H., Vecht U., Gielkens A.L.J., Smits M.A.;
RA	"Repeats in an extracellular protein of weakly pathogenic strains of
RF	Streptococcus suis type 2 are absent in pathogenic strains.";
RL	Infect. Immun. 61:3318-3326(1993).
DR	EMBL: X71880; CAA50714.1; .
DR	PIR: S33441; S33441.
DR	GO: GO:0016020; C:membrane; IEA.
DR	InterPro: IPR005877; Gpos_Y5IRK.
DR	Pfam: PF04650; Y5IRK signal; 1.
DR	TIGRFAMs: TIGR01168; Y5IRK signal; 1.
SQ	SEQUENCE 1822 AA; 192632 MW; 3838960C77641D7D CRC64;

Query Match	7.2%;	Score 205;	DB 2;	Length 1822;
Best Local Similarity	23.7%;	Pred. No. 0.3;		
Matches 145;	Conservative 82;	Mismatches 229;	Indels 156;	Gaps 27;

QY	37	PTFDYKTAQ-----TAYDTFTSTSLADIQAALVSLQDA--VTNIKDTAA	81
Db	1265	PNLSDEKQAEIKKLTDAVAKTATLRNADKKTQAEKAQALADLEKAKETQKIADKAA	1324
QY	82	TDEETAIA--AEWETKNADAVKVGAGITELAKYASDNQAILDSLGLTSPDLLQALLOS	139
Db	1325	IDRLTILVKDGELEATKQDAKNKIADAAAKEAIAASNPNLTDAEKTFDVAIDA--E	1380
QY	140	VANNKA-----AELLKEMQNPVVGKTPAIAQSLVDQTATA-TQTEKDGNAIRDA	191
Db	1381	VAKANDAISAATSPADVQKE-EDAGVA-----AIAEDVLDAAKQDAKNKIADAAAKEA	1434
QY	192	YFAGQNASGA-----VENAKNNSNTSDSAKAALATATAKTQIAEAAQKFPDSPIL	241
Db	1435	IGSNPNLTDAEKTFDVAIDAEVAKANDALS-----AATSPADVQKE-EDAGVA	1482
QY	242	QEAQWVIAQEKDLKNIKPADGSDVPNPGTTVGSGKQGGSSIGSIRVSMILLDDAENETAS	301
Db	1483	AIAEDVLDAKQDAKN-KIAKESD-----AAKSAIDANPNLTDAEKESAK	1526
QY	302	ILMSGFRQIMHNTENPSSQAQELAA-----QARAKAAGDSAAALADAQK----	352
Db	1527	-----KAVDADAKAATDAIDASTSPVEAQSABDKGVGSIQADVLDAAKQDAK	1573
QY	353	---ALEAALGKA-----CQQQGIILNALGQIASAA--VVSAGVPPAAASS-----I	392
Db	1574	NKTAKEVAAAKEAIDANPNLSDEKESAKKAVDADAKATTDAIDASTSPVEAQSABDKGV	1633
QY	393	GSSVKOLYTKSTGSDYKTIQISAGVDAYKSIINDAYGR---ARNDATRDVIN---NVSTP	446
Db	1634	GSIRQDVLDAAK--QDAKNKIAKESDAKSAIDANPNLTDAEKESAKKAVDADAKAATD	1690
QY	447	ALTRSVPRARTKARGPE-KTDQALRVI-----SGNSRTLGVIVYSGVSALQSMOITQS	499
Db	1691	AIDAST--SPVEAQSABDKGVGAIAKDILDAKQDAKNKIADAAKEASAKSVIDSNPNLTD	1748
QY	500	NPOANNEIRQKLTSAVT-----KPPQFCGYPVYQLSNDSTQKF	537
Db	1749	AKEGAKESEIDKAVEEAIVLINGVRTYQEELEKILPMAALIKPAAKTPVVD-FNNLTEXE	1807
QY	538	IAKLESLFAEGS	549
Db	1808	IARIKAFLENN	1819

RESULT	9
--------	---

RESULT 9


```

Db 2663 GNYTNASDCKGAYTDAYNAAKNIIVGSPNVTINRAADVTAATQVRVNNNAETSLNGDTNLAT 2722
Qy 462 -PEKTDQALARV-----ISGNSRTLGDVYSQVSALQSVNQITQS-----NP 501
Db 2723 AKQQAQALRQWTHLSDAKQSITQI-----DSATQVTGVQSVKDNATNLNANQOLNS 2778
Qy 502 QANNEIRKQUTSVAVKPPQPGYFYVQLSNDSTOKF-----IAKLSLFAEGSRTAAEIKAL 558
Db 2779 IANKDEVKA-----SQPYVDADTDKQAYNTAVTSAENIINATSQPTLDPSAV 2826
Qy 559 SFETNSLFIQVILVN 573
Db 2827 TQANQVNTKALN 2841

RESULT 11
Q8CP76 PRELIMINARY; PRT; 9439 AA.
AC Q8CP76;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Eha protein.
GN SE1128
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016747; AAC04725.1; -.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR002988; GA.
DR Pfam; PF05738; Cna_B; 1.
DR Pfam; PF01468; GA; 39.
KW Complete proteome.
SQ SEQUENCE 9439 AA; 1050771 MW; 5A8C68E9AA31CDD4 CRC64;

Query Match
Best Local Similarity 7.0%; Score 200.5; DB 16; Length 9439;
Matches 132; Conservative 99; Mismatches 255; Indels 161; Gaps 27;

Qy 12 SNSSSTSRSDVSTTATATPTPPPTFDY-----KTQAQTAYDTFTSTSLAD 61
Db 1945 NNGDATSQINDAKNTVEQAKR-----DYVEAKSNLRADKSQLQASDYTLNRDVLTD 1997
Qy 62 IQAALV-SLQDAVNIKDTATDEETATAAEWETKNADAVKVGQITELAKYSDNQAIL 120
Db 1998 KKPASVRRYNEAISNIR-----KELDTAKADASST-LRNTNPSVEQVR 2039
Qy 121 DSLGKLTSD--LLQA--ALLOSVAANNKAAELIKEMQDN-----PVVPGKTP----- 164
Db 2040 DALANKINTVQKVNQAIALLPKENNSLVQAKKRLQDAVNDIQTQGMTOQTINNYNDK 2099
Qy 165 -----ALAQSLVDQTDATATQIEKDGNAIRDAYFAGQNASG-----A 201
Db 2100 QREABRALTSAGRVINDGDAITQETITSEKSKVEQAMQALTNAKSNLRADKNELOTAYNKL 2159
Qy 202 VENAKSN--SISNIDSAKAAIATAKTOIAEAQKFPDPSILOAEQMWIAEKDLKN 257
Db 2160 IENVSTNGKKPASTROYETAKARI---QNQINDAKNE--AERILGNDNPQVSQVTOALNK 2214
Qy 258 IKPADGSDVPNPGTIVGSGSQGSSIGSIRVSMILDDDAENETASILMSGFQMIHMENTE 317
Db 2215 IKAIC---PKLTEAINLMQKENNELVNAKRLNENAVNTDTP--THGMTQ--ETINNY 2266
Qy 318 NPDSQAQOEIAAQARAAKAGDDSAALADAQKALEAALGKAGQOQGIILNALGQIASA 377
Db 2267 NAKKREAQNEI--QKANMIINNGDATAQDISSEKSKVEQVL-----QALQNAKNDLR-- 2316

```

```

Qy 378 AVWSAGVPPAAASIGSSVQLYKXTSKTG-----SDYKTIQISAGYDAYKSIINDAYGRA 431
Db 2317 -----ADKRELQAYNKLQNVNTNGKPPSIQNYKS-----ARRNIENQYNTA 2360
Qy 432 RNDA-----TRDVIN-----NVSTPALFESVPRARTE-----ARGPEKTDQALA 470
Db 2361 KNEAHNVLENTPTVNAVEDALRKINAIQPEVTKAINILOKEDNSSELVRAKEKLDQAIN 2420
Qy 471 RVISGNSRTLGDVYSQVSALQSVNQITQS-----NFOANNEEIR-----QKLTSAVTK 518
Db 2421 SQPSLNCMTQESINNYTTKREAQNIASSADTIINNGDASIEQITENKIRVEBATNALNE 2480
Qy 519 PPOFGYPYVQLSNDSTOKFAKJESLFAEGSRTAAEIKALSFETNSL 565
Db 2481 AKQ-----HLTADTTSLKTEVRKLRRGDTNNKKPSSVSAYNNTIHSL 2523

RESULT 12
Q9NE28 PRELIMINARY; PRT; 1313 AA.
AC Q9NE28;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN L232.01.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Klages S., Borzym K., Beck A., Reinhardt R., Ivens A.C., Quail M.,
Rajandream M.A., Barrell B.G.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL; ALJ54553; CAB89628.2; -.
KW Hypothetical protein.
SQ SEQUENCE 1313 AA; 140188 MW; 2BA3034A75AD23AC CRC64;

Query Match
Best Local Similarity 7.0%; Score 200; DB 5; Length 1313;
Matches 126; Conservative 97; Mismatches 251; Indels 88; Gaps 25;

Qy 28 TATATPTPPPTFDYKTAQTAFTDTFTSTSLADIQAA-LVSLQDAVNTIKDTAATDEET 86
Db 244 SSTTPTPPSS-----DMAYQLTLEBSTAAHFEDLQRRITAVVASSRDAFV 291
Qy 87 AIAAEWETKNADAVKVGQITELAKYASDNQAILDSLGLKLTSPDLLQALLOSVAANNKA 146
Db 292 SAPAELETAVVEAVQ-----KAAAEQ-----EKLNLINDLLAEERLAARGSGGDA 339
Qy 147 AEILKEKQDNPNVPGKTPAIQAQSLVDQTDATAT-----QIEKDGNAIRDAYFAGQ--NASG 200
Db 340 ANAIVDMR-----ATEESTESAVQSATNELRTITSEVKKLEK-SKGTGEVTAGSATDAM 395
Qy 201 AVENAKSNNSISNIDSAKAAIATAKTOIAEAQKFPDS-PILOAEQMWIAEKD-LKNI 258
Db 396 AIEIEKVFDAVQSQSLAAVDEVVGGDFKEQAATRDALRVQDE----VLQRLDELURGT 451
Qy 259 KPADG-SDVNPGETT-VGSGKQOQSSIGSIRVSMILDDDAENETASILMSGFQMIHMENT 316
Db 452 VATEATASTPSPAAPYNGVPEAAQQ---KITEVLEQLRRD-----VSATSSMLEIHN 502
Qy 317 ENPDSQAQOEIAAQARAAKAGDDSAALADAQKALEAALGKAGQOQGIILNALGQIAS 376

```

```

Db      503 N-----AQOQLMT---VAQWSDSVTAALQOHTTEALQSAMTTVOFQLNALREAGEAAS 553
QY      377 AAVVSAGVPPAAASSGSSVKQLYKTSKSGDYKTOISAGYDAYKSINDAYGRANDAT 436
Db      554 AAASAAPPAAAPSPAPLSOEVLTDVAEAVLSRPAWELSV---QVEKNTANKLALIEQQL 610
QY      437 ROVINNVST---PALTRSVPRARTEARGPEKTKQALARVISGNSRSLGDIYSQVS--ALQ 491
Db      611 LHIADSVSTVAASLOAQLEEMRTSABESQLKTEHA-----AKEGVSVIDTSPAATA 662
QY      492 SVMQITQSNPOANNEIROKLTSAVTKPPQFG-----YPYVQLSNDSTQKFIKLESIFAE 547
Db      663 ALIEKTHEATTREELQLAQVQKTPENVLTMPYKYI---DSVLTPMK--EELSLO 716
QY      548 GSRTAAEIKALSPETNSLFIQ 569
Db      717 ETNLTKAMASLA-ETVTAAVRE 737

RESULT 13
Q8NWQ6
ID Q8NWQ6 PRELIMINARY; PRT; 9904 AA.
AC Q8NWQ6
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Ebn protein.
GN EBN OR MW1324.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004826; BAB95189.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002988; GA. YSIRK.
DR InterPro; IPR005877; Gpos_YSIRK.
DR InterPro; IPR006530; YD.
DR Pfam; PF01468; GA; 48.
DR Pfam; PF04650; YSIRK signal; 1.
DR TIGRFAMs; TIGR01643; YD repeat 2x; 8.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
KW Complete proteome.
SQ SEQUENCE 9904 AA; 1068476 MW; CB86908F75EA5682 CRC64;

Query Match 7.0%; Score 200; DB 16; Length 9904;
Best Local Similarity 20.8%; Pred. No. 4.6;
Matches 143; Conservative 103; Mismatches 262; Indels 180; Gaps 29;

QY 5 DKIGIASSNSSSTSRGADVSTTATPTPPFDYDKTQAQTAYDTFTSTSLADIQA 64
Db 5483 EKLAARQAKTSKDITGRULDLNNAQRTA-----ANAEVDQAPLAAVTA 5525
QY 65 A---LVSLQDAVTNKKTAATDEETAIAAEWETKNADAVKVGAGITELAKYASDNQAILD 121
Db 5526 AKNKATSLNTAMGNLKHAAEKONTKRSVNY--TDADQPKQQAQVDTAV---TQAEALTN 5579
QY 122 SLGKLTSPDLQALLO-----SVANNKKAELLKEM 153
Db 5580 ANGSAENETTVOALNQLNKAQNDLNGDNKVAQAKESAKALASYNLNAQAQSTAAATSOI 5639
QY 154 QDNVPVPGKFTAIQSLVDQDTATATQIE---KGNAI-----RDAYF-AGQ 196
Db 5640 DNATTVAGVT--AAQNTANELNTAMGQLQNGINDQNTVKQVNFDTADQGGKDAYTNVAT 5697
QY 197 NASGAVENAKNSNISIDSAAIAATAKTQIAEAKQKFPDPSILQEAQWVQAEKD-- 254

```

```

Db      5698 NAQGLDKANGQN-----MTKAQVEAALNQVTTAKNAL-----NGDANVRQAKSDAK 5744
QY      255 -----LKNIKPADGSDVPNPPTVGG-----SKQOG-----SSIGS---IRVS 289
Db      5745 ANLGLTTHLNNACKQLDTSQIEGATTVNGVSVKTKAQLDGMORLESIAINKDQTKAS 5804
QY      290 MLDDDAENETASILMSGFRQMIHMENTE---NPDSQAQQLAAQAAAKAA--GDSAAA 345
Db      5805 ENVIDADPTKKTAFDNATQAESYLNKDHGTNKKQAVQQAICSVTSTENALNGDANLQR 5864
QY      346 ALADAQKALEALGKAGQOQGITNALGQIASAAVVSAGVPPA--AASSTGSSVKQLYKTS 403
Db      5865 AKTEATQATDNLTHLNTPKQ--TALKQOVNAAQRVSGVTDLKNATSATSLDNAMDQL--- 5917
QY      404 KSTGSDYKTCISAGY-----DAYKSINDAYGRARN--DATRDVINNVV--TPALTRESVPR 454
Db      5918 KQGIADHDITVAGGNVTNAPSKQAGYTDAYNAAKNIVNGSENVITNAADVTAATQRVNN 5977
QY      455 ARTEARG-----PEKTDQALARV-----ISGNSRTLGDVYSQVSALQSVMQIT 497
Db      5978 AETSLNGDSNLATAKQAKDALRQWTHLSDAQKQSITGOI----DSATQVTGVQSVKDNA 6033
QY      498 QS-----NPOANNEIROKLTSAVTKPPQFGYPYVQLSNDSTQKF---IAKLESIF 545
Db      6034 TNLDNAMNQLRNSIAKDEVKA-----SQPYVDADTDKQNAVNTAVTSAENII 6081
QY      546 AEGSRTAABEIKALSPETNSLFIQOVLVN 573
Db      6082 NATSQTLDPASVTDQANQVNTKNTALN 6109

RESULT 14
Q8NUU3
ID Q8NUU3 PRELIMINARY; PRT; 2275 AA.
AC Q8NUU3
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE MW2575 protein.
GN MW2575.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AP004831; BAB96440.1; -.
DR InterPro; IPR008985; ConA_like_lec_gl.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR008009; He_PIG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF05345; He_PIG; 2.
DR PROSITE; P050847; GRAM_POS_ANCHORING; 1.
KW Complete proteome.
SQ SEQUENCE 2275 AA; 228407 MW; E676B6BB9B60CE2E CRC64;

Query Match 7.0%; Score 199.5; DB 16; Length 2275;
Best Local Similarity 19.3%; Pred. No. 0.73;
Matches 115; Conservative 127; Mismatches 235; Indels 119; Gaps 22;

QY 2 SLAD---KLGIASSNSSSTSRGADVSTTATPTPPFDYDKTQAQTAYDTFTSTLS 58
Db 1236 SLSDSTSESGSTSLSNSTSGSASISTSTISEST-----STFKSESVSTSLSMSTSTS 1290
QY 59 LADIQAALVSLQDAVTNKK--DTAATDEETAIAAEWETKNADAVKVGAGIT-ELAKYASDN 116

```



```
Db 1291 LSDSTSLSTSDSTSDKSDSLSTSMSTSDSI--STKSDSISSTSLSTSGSTSESDS 1348
Qy 117 QAILDLSGLKLSFOLLQAALLQSVANNKKAELLKEMQNPVPGKTPAIAQSLVDQIDA 176
Db 1349 TSSSEKSDSTSMISI---SMGSTSGSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1384
Qy 177 TATQTEKDNARIDAYFAGQNASGAVENAKSNNSISNIDSAKAAIATAKTOIAEAKKFP 236
Db 1385 TSTSL-----SLSAMNQSG-VDGNSAQSSASN-----STSTSTSTSDSQST 1425
Qy 237 DSPILQAEQVIOAEKDLKNIPKADGSDVPNPGFTVGGSKQGSIGSIRVSMLLDAAE 296
Db 1426 SG-----YTSQSTQSESTSTSTSLSD-----STSIKSTSGSGSV-STASLSTSGSSE 1473
Qy 297 NETASILSGFQMIHMFNTENPDQAAQCELAQA-----ARAAGAAGDGAALADAQ 351
Db 1474 SDQSISSTASSTSESASTSLSDSTSTSNSSGASTSTSLNSASASSDSSSTSLSDST 1533
Qy 352 KA-LEALGKAGQQQGI-----NALGQIAAAVVAGVPPAAASSIGSSVKQLYK 401
Db 1534 SASMQSSSDSQSTGASLSDSTSTSNMSTIAS-----LSTSVSTSESGSTSESDS 1590
Qy 402 TSKSTGSDYKT--QISAGVDAYSINDAYGRANDAT---RDVINNVSTPALTSVPRA 455
Db 1591 TSTSLSDSQSTGRSTASGASASTSTSTSDRSTASTSTMTSTSTSDSQMSLSTSTST 1650
Qy 456 RTEARG-----PEKTDQALARVISGN---SRTLGVDVYSQVSLQSVMQITQNPQANNE 507
Db 1651 MSDSTSLSDSVSDSTSDSTASGSMVSISLSDSTSTSTSAFVMSASISDSQSMSES 1710
Qy 508 IRQKITSVTKPPQPGYPIVQLSNSTQKFIKLSLFAEGRTAEIKALSFETN 563
Db 1711 VN-----DSRSVSESNESDSKSMGSGTSTVSDSGSLSVSTS 1746

RESULT 15
Q8XDQ4
ID Q8XDQ4 PRELIMINARY; PRT; 973 AA.
AC Q8XDQ4;
DC Q8XDQ4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative membrane protein of prophage CP-933X (Putative tail fiber protein).
DE GN Z1918 OR ECS1650.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
RL EMBL; AB005333; AAG56007.1; --
```

```
DR EMBL; AP002555; BABJ5073.1; ALT_INIT.
DR PIR; B90835; B90835.
DR PIR; C85693; C85693.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR008969; CarboxypepD_reg.
DR InterPro; IPR005003; Phage_fiber.
DR InterPro; IPR005068; Phage_fiber_2.
DR Pfam; PF03335; Phage_fiber; 6.
DR Pfam; PF03406; Phage_fiber_2; 3.
KW Complete proteome.
SQ SEQUENCE 973 AA; 96316 MW; 776580D2A87E1B36 CRC64;

Query Match 7.0%; Score 199; DB 16; Length 973;
Best Local Similarity 20.6%; Pred. No. 0.26;
Matches 133; Conservative 79; Mismatches 227; Indels 206; Gaps 20;

Qy 9 IASNSSSSTSRSDVD-----STTATAPTPPP-----TFDD 41
Db 40 VASENPDEAGRYSDVEYQYSVILLVEGFPSPHAGTITVYEDSQPTLNDFLGAWTEDD 99
Qy 42 YKQATQATYDTIF-----TSTSLADIQAALVLSLQDAVTNIKDTAATDEETAIAAEWTK 95
Db 100 VEPEALRRFELMVEEVARNASAVAQNTAAAKKSADASTSARE-----AAATHATD 149
Qy 96 NADAVKVGQAQITELAKYASDNQAILDSLGLKLTSPDLLQAALLQSVANNKKAELLKEMQD 155
Db 150 AADSAR--AASTSAGQAASSAQSASSAGTAST----- 180
Qy 156 NPVVGKTPAIAQSLVDQTATQIEKDGNAIRDAYFAGQNASGAVENAKSNNSISNID 215
Db 181 -----KATEASKSAAAEASSKSAATSAAGAKTSETNAVSQ-Q 218
Qy 216 SAKAAIATAKTQIAEPAQKFPDPIQAEQVIOAEKDLKNIKPADGSDVPNPOTTVGG 275
Db 219 SAATSASATTTKASEAASSARDASAKEA-----AKGSETSAASSASSAASS 265
Qy 276 SKQGSSTIGSIRVSMLLDDAENETASILMSGFRQMIHMFNTENPDQAAQCELAQAQAAA 335
Db 266 ATAAGNSAKAAKTS-----ETNAKSETAAAEQSAS 295
Qy 336 KAAGDSDSAAALADAQAALGKAGQQQGIHALGQIAAAVVVSAGV-----P 385
Db 296 AAGSKTAAALSA-----SAASTSAGQASASATAAGKSABSAASSASTATTKAGEATEQ 349
Qy 386 PAAASSIGSSVKQL---YKTSKSTGSDYKTOISAGYDAYKSINDAYGRANDATR--DVI 440
Db 350 ASAAASSAAAKTSETNAKASETSAESSKTAASASSASSASSASSASASAKDEATQASAA 409
Qy 441 NNVSTPALTSVPRA-----RTEARGPEKTDQALARVISGNSRTLGD 482
Db 410 KESATTAATKATEAAGSATAAAQSKSTAESAATRAETAAKRAEDIASVALEDASTTKKG 469
Qy 483 VYSQVSALQSNMQITQSNPQ-----ANNEEIRQKLTSAVTKPPQFGYPVVLSEN--D 532
Db 470 IVQLSSAINTSSESIAATPKAVKAAYELANGKYTAQDATT-----QXG--IVQLSNATN 522
Qy 533 STQKFI-----KLESIFASGSRTPAE-----IKALSFTNS 564
Db 523 STSEMLAATPKSVKAAAYDLANGKYTAQDATTQAQKIGVQLSSATNS 567

Search completed: March 24, 2004, 05:59:12
Job time : 46.9152 secs
```